An Improved Model of Wouthuysen-Field Coupling around Galaxies

Mikael Toresen

Thesis submitted for the degree of
Master of Science in Astronomy Master of Science

Institute of Theoretical Astrophysics
University of Oslo

June 1, 2015

Acknowledgements

I would like to thank my thesis supervisor, Mark Dijkstra, and the extragalaactic astrophysics group constructive discussion, support and encouragement. I would also like to thank Jonathan R. Pritchard for graciously shared his code for computing the Lyman decay probabilities and recycle fractions, allowing me to verify that deviations between my results and previous ones were not due to errors.

Table of Contents

Preface

In this thesis We will look at the effects of the Wouthuysen-Field coupling of Lyα on the IGM surrounding a central source at $z = 25$ with an ionized bubble, powerlaw UV spectrum, atomic and molecular hydrogen absorption features, when assuming wing scattering affects the radial distribution of photons in the thermal core.

We will first go through the physics describing the universe during the dark ages and reionisation of the universe, and the structure formation which culminates in galaxy and star formation. Then we will look at the 21cm hyperfine transition of the hydrogen atom, before describing the model and its implications.

Throughout the thesis we will attempt to describe the causality chain which leads to star formation, but discussing global distributions like correlation functions, IMF, luminosity functions, power spectra is beyond the scope of this thesis. We have also tried to keep to avoided showing best fit values from models with numerical estimates, unless directly applicable to the model, and we try to emphasise the underlying physics.

Throughout the thesis we have tried to stick to SI units for numerical values, so the reader with experience in this field using cgs is cautioned to keep this in mind. The formulae, however, are all in cgs, as is customary in this field.

Chapter 1

Background

1.1 The Friedman equations

In 1915, Einstein presented the theory of general relativity, which became the framework upon which our modern understanding of cosmology and our universe. One particularily important set of equations for cosmology that arises as a solution to Einsteins field equations is the *Friedmann equations*. These describe a homogeneous and isotropic universe which is either expanding or contracting. In this description the universe is filled with a perfect fluid with a given energy density ρ and pressure p, dependent upon the content of the universe:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = \frac{8\pi G}{3}\rho \tag{1.1}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \tag{1.2}$$

Here $a(t)$ is the scale factor, which parameterizes the size of the horizon, and k is the curvature of the universe at $a = 1$ (today). "$\dot{}$" denotes the time derivative $\partial/\partial t$. Current observations indicate that $k = 0$, which corresponds to a flat (Euclidean) universe, or alternatively, as will see setting $\Omega \equiv \frac{\rho_{\mathrm{tot},0}}{\rho_{\mathrm{cr},0}} = 1$. Where $\rho_{\mathrm{cr}} \equiv \frac{3H^2}{8\pi G}$, $H \equiv \dot{a}/a$ is the Hubble parameter and the subscript 0 implies values today.

While this is a very simple description of the universe, it provides a good understanding of how we understand the universe on its largest scales.

The components of the universe evolve differently with time depending on their equation of state, or the relation between (energy) density and pressure. For the cosmological background, this is described using the simple form $p = w\rho c^2$, where w is in most cases assumed to be constant. Assuming adiabatic expansion we can find an expression for the evolution of the energy density:

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \tag{1.3}$$

By sorting the equation then integrating from a given energy density to the energy density today $\rho_0 = \rho(a = 1)$ we get:

$$\rho = \rho_0 a^{-3(1+w)} \tag{1.4}$$

In modern cosmology several components are considered which can be roughly divided into the following three categories:

- **Radiation,** $w = \frac{1}{3}$**:**This includes all massless particles and to a good approximation particles with energy much greater than its rest mass. Photons and neutrinos are treated as such.
- **Dust,** $w = 0$**:** Non-relativistic pressureless matter. This is usually divided into two separate constituents: baryonic matter, which we know well from the periodic table, and dark matter, which does not interact much with baryons or photons. For our purposes, the main thing we need to know about dark matter is that it behaves like collisionless particles; The only interact through gravitation(and possibly also though weak interaction).
- **Cosmological constant,** $w = -1$**:** Dark energy. From (1.6) we see that with $w = -1$ the energy density is constant. Little else is known about this other than it is necessary to explain the accellerated expansion of the universe.

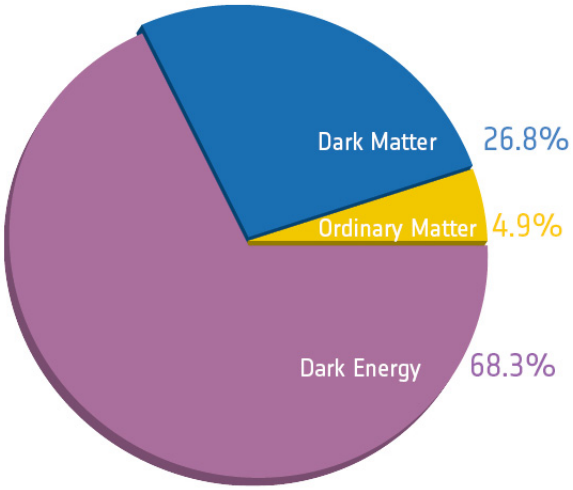

Figure 1.1: Relative material content of the universe. Here "ordinary matter corresponds to baryonic matter. *Image credit: ESA/Planck*

Each of these components can, to first order, be considered as independent of eachother. One can therefore separate out each of the components of ρ which can also be rewritten in

terms of their values today using (1.6):

$$\rho = \rho_r + \rho_b + \rho_m + \rho_\Gamma = \rho_{r,0}a^{-4} + \rho_{b,0}a^{-3} + \rho_{m,0}a^{-3} + \rho_{\Gamma,0} \tag{1.5}$$

Furthermore we can relate this to the critical density today:

$$\begin{aligned}\rho &= \frac{3H_0^2}{8\pi G}\frac{1}{\rho_{\mathrm{cr},0}}\left(\rho_{r,0}a^{-4} + \rho_{b,0}a^{-3} + \rho_{m,0}a^{-3} + \rho_{\Gamma,0}\right) \\ &\equiv \frac{3H_0^2}{8\pi G}\left(\Omega_{r,0}a^{-4} + \Omega_{b,0}a^{-3} + \Omega_{m,0}a^{-3} + \Omega_{\Gamma,0}\right)\end{aligned} \tag{1.6}$$

Where $\Omega_i \equiv \rho_{i,0}/\rho_{\mathrm{cr},0}$. Inserting this into (1.1), we can rewrite the FRW equation:

$$\begin{aligned}H^2 + \frac{kc^2}{a^2} &= \frac{8\pi G}{3}\frac{3H_0^2}{8\pi G}\left(\Omega_{r,0}a^{-4} + \Omega_{b,0}a^{-3} + \Omega_{m,0}a^{-3} + \Omega_{\Gamma,0}\right) \\ \frac{H^2}{H_0^2} + \frac{kc^2}{a^2H_0^2} &= \Omega_{r,0}a^{-4} + \Omega_{b,0}a^{-3} + \Omega_{m,0}a^{-3} + \Omega_{\Gamma,0}\end{aligned}$$

Assuming $\sum \Omega_{i,0} \equiv \Omega = 1$ and looking at $a = 1$ we see that this corresponds to a flat universe, $k = 0$. We then get our final expression:

$$\frac{H^2}{H_0^2} = \Omega_{r,0}a^{-4} + \Omega_{b,0}a^{-3} + \Omega_{m,0}a^{-3} + \Omega_{\Gamma,0} \tag{1.7}$$

From (1.7)we see that the expansion of space is dependent on all the different densities, which can make this equation difficult to deal with. A common simplification when looking at the universe around a given redshift, is to only look at the main components. When $\Omega_i a^{-3(1+w_i)} \gg \sum_{j\neq i}\Omega_j a^{-3(1+w_j)}$ we say that the universe is dominated by that component. Thus at very early times we need only look at the radiation term(*radiation-dominated universe*), then at later times dust becomes important(*matter-dominated universe*), and so on.

1.2 A brief overview of the early cosmic history

Now that we have a simple picture of how the universe evolves we can explore the stages the universe has undergone from the Big Bang.

Current physics break down on Planck scales, where the age of the universe is on the order of $t_{\mathrm{Pl}} \approx 10^{-43}$ s. At this time the universe is of the size of $l_{\mathrm{Pl}} \approx 10^{-35}$ m. At these very early times quantum effects dominate, and we expect there to be some quantum fluctuations in this initial hot plasma. In the midst of this (at some time after t_{Pl} believed to be around $t \approx 10^{-30}$ s.) we believe the universe underwent a rapid expansion called *inflation*. Inflation causes any initial curvature to flatten out, and causes regions which were causally connected to become much larger than otherwise. Due to quantum fluctuations

we do not expect inflation to stop simulaneously everywhere. In fact, we expect that due to these flucutations the universe will become inhomogeneous enough to eventually form the large-scale structure we see today [12].

Some time after this we expect that dark matter will no longer interact with baryonic matter in any other way than with gravitation. Dark matter this will naturally be pulled towards local overdensities, which in turn leads to the matter distribution becoming further polarized. This is often described as a potential well, becoming deeper as more matter "falls" into it. This happens over a very large span of time-scales, with clumping first starting on small scales then later on large scales.

When the age of the universe is of the order of 1 second we expect protons and neutrons to form, while still being unable to form atoms, as they immediately get ionised by the high density of photons. These in turn form deuterium, and then helium as the temperature decreases. Once it becomes cold enough, this reaction stops, and the ratio of hydrogen and helium in the universe today is largely decided by how long this reaction could happen.

After $7 \cdot 10^4$ years we expect matter to start dominating the expansion of the universe. The pressure from the photons is now gradually diminished due to the expansion and lowered temperature, and baryonic matter now also starts falling into the potential wells created by the quantum fluctuations and amplified by the dark matter.

At around $4 \cdot 10^5$ years after the Big Bang, the universe has is cold enough for atoms to form, effectively making the universe neutral. Photons(which are the force carriers of electromagnetism) are thus able to travel freely instead of being scattered off ions. Thus it is often referred to as the *last surface scattering* and is the origin of what cosmologists refer to as the *cosmic microwave background* (CMB), which is central to our understanding of the early universe.

After the last scattering the universe enters what is called the dark ages, where photons interact minimally with baryons, making this epoch very difficult to observe. Approximately 10^6 years after the Big Bang, the density of mass in the largest potential wells has caused the first stars and galaxies to form. Once these primordial stars start radiating photons, the gas nearby will in turn ionise. Eventually this occurs everywhere in the universe, such that the universe once again is ionised. This is referred to as *reionisation*, after which the modern universe emerges.

1.2.1 Temperature evolution

To determine the evolution of the universe, one essential quantity we need to know is the temperature. The temperature is what allows us to determine the population of atoms and molecules at a given time, and thus when we can expect stars and galaxies can form. In addition, as will be explained in 3, knowledge of the average temperature of the CMB at a given redshift will be essential to whether and how the effects of the spin temperature will be observed. To determine the evolution of the average temperature we first solve the equation for the evolution of the energy density of radiation, which is closely linked to

temperature. Rewriting (1.3) we get:

$$a^{-3}\frac{\partial\left(\rho a^3\right)}{\partial t} = -3\frac{\dot{a}}{a}\frac{p}{c^2},$$

which is the energy conservation law in an expanding universe. Inserting $p = \rho c^2/3$, $\rho = \rho_r$, we get:

$$\begin{aligned} a^{-3}\left(\dot{\rho}_r + \frac{\partial a^3}{\partial t}\right) &= \frac{\dot{a}}{a}\rho_r \\ a^{-3}\dot{\rho}_r + 4\frac{\dot{a}}{a}\rho_r &= 0 \\ a^{-4}\frac{\partial \rho_r a^4}{\partial t} &= 0 \end{aligned}$$

implying $\rho_r \propto a^{-4}$. Using the knowledge that photons follow the Bose-Einstein energy state distribution in equilibrium at a given temperature, we can also write:

$$\rho_\gamma = \frac{g_\gamma}{(2\pi)^3}\int_0^\infty \frac{1}{e^{E(p)/kT}-1} p\, d^3p = \frac{\pi^2}{15}T^4, \tag{1.8}$$

where p here is momentum, and the chemical potential of the Bose-Einstein distribution is zero(due to photon number not being conserved) and the degeneracy of photons in each energy state is $g_\gamma = 2$. By setting $\rho_r = \rho_\gamma$ we then arrive at our relation between temperature and redshift:

$$T \propto \rho_r^{1/4} \propto a^{-1} = (1+z) \tag{1.9}$$

$$T = T_0(1+z) \tag{1.10}$$

Note that this is a zero-order approximation and only applies to the very early universe due to the assumption of thermal equilibrium. It does however give a very good idea of the average temperature of the CMB. The only thing left to do is to find the proportionality constant T_0. This can be found in the CMB-temperature today, which is measured to be approximately $T_0 = 2.7$ K.

1.3 Structure formation

In this section we will attempt to give a simple overview of how large scale structure evolve in the linear regime, how particles interact to form the overdensities in which galaxy clusters could form. Note that several simplifications are made; for readers looking for a more rigorous review, we would recommend [4, 40].

1.3.1 Newtonian linear perturbation theory

In section 1.2 we briefly hinted that overdensities are formed by the gradual evolution of what was initially only quantum fluctuations, differences in densities on the planck scale. If we in general describe deviations from the average value by δ given as:

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho(t)}}$$

where $\bar{\rho}$ is the average density at a given time given by (1.6), then we can get insight into how these deviations evolve using linear pertubation theory. This assumes we can write all quantities as slight deviations from the average value ie. $|\delta| \ll 1$, or more generally:

$$X = X_0 + \Delta X, \quad \text{where } X \in [\rho, p, \phi, \mathbf{v}], \quad \frac{\Delta X}{X_0} \ll 1.$$

This is a good approximation at early times and on very large scales, in what is referred to as the *linear regime*. We then want to find the equations of motion caused by gravitation and pressure. Expressing physical quantities using comoving units:

$$\mathbf{x}(t) = a(t)\mathbf{r}(t) \tag{1.11}$$

$$\Delta\mathbf{v}(t) = a(t)\mathbf{u}(t) \quad \Rightarrow \quad \dot{\mathbf{x}} = \dot{a}\mathbf{r} + a\dot{\mathbf{r}} = H\mathbf{x} + \Delta\mathbf{v} \tag{1.12}$$

Note that $\mathbf{v}$ is related to the peculiar velocity, and $H\mathbf{x}$ to the Hubble flow. We then linearize the following equations:

$$\frac{D\rho}{Dt} = 0 \tag{1.13}$$

$$\rho\frac{D\mathbf{v}}{Dt} = -\nabla p - \nabla\phi \tag{1.14}$$

$$\nabla^2\phi = 4\pi G\rho \tag{1.15}$$

Where $D/Dt \equiv \partial/\partial t + \mathbf{v}\cdot\nabla$, and ϕ is a gravitational potential. To solve these equations, we Fourier transform the density perturbation:

$$\delta(\mathbf{r}, t) = \sum_k \delta(\mathbf{k}, t) e^{i\mathbf{k}\cdot\mathbf{r}},$$

where $\mathbf{k}$ is the comoving wavevector. Following eg. [40] we get the equation of motion for the density perturbation:

$$\ddot{\delta} + 2H\dot{\delta} = \delta\left(4\pi G\rho_0 - \frac{c_s^2 k^2}{a^2}\right) \tag{1.16}$$

where $c_s = \frac{\Delta p}{\Delta \rho}$ is the speed of sound, and the 0 subscript here indicates the linearized expression. How can we interpret (1.16)? If there is no Hubble expansion, δ is determined by the perturbation of the gravitational potential $\Delta\phi$ and the pressure p. If there is a large gravitational potential perturbation, an overdensity will continue to evolve, unless the pressure gradient that is developed hinders it from doing so. Pressure will attempt to bring the perturbations towards 0. The Hubble expansion streches all scales and this also effectively acts as a pressure on overdensities on very large scales, but also acts as a negative pressure on underdensities giving rise to voids.

1.3.2 Jeans scale

The interplay between the pressure gradient and gravity described in the end of subsection 1.3.1 is central to determining if large structures form. We therefore would like to have a measure of the relative strengths between these two forces. One such quantity is the *Jeans length* λ_J, which is given as the wavelength scale at which these two forces are equal.

$$4\pi G\rho_0 = \frac{c_s^2 k_J^2}{a^2}$$
$$\lambda_J = \left(\frac{2\pi}{ak_J}\right) = c_s\sqrt{\frac{\pi}{G\rho}}$$

Where in the last equality we set $\rho_0 \approx \rho$. Similarily we can define the *Jeans mass* as the matter contained within a sphere of radius $\lambda_J/2$.

$$M_J \equiv \frac{\pi}{6}\lambda_J^3\rho \tag{1.17}$$

If the Jeans mass within this sphere $m > M_J$ then we expect gravitational collapse, while for lower values pressure prevents this.

Revisiting (1.16), what kind of solutions to δ do we expect? This equation is simply the equation for damped oscillations with a negative friction term. Assuming $H = 0$, when $\lambda > \lambda_J$ we expect an exponential solution, which cannot be interpreted as anything but gravitational collapse. A more detailed calculation taking into account the expansion of the universe the solution will give growth due to gravitational collapse as a power law [31]. For smaller scales $\lambda < \lambda_J$ we see that this turns into a harmonic oscillator, giving rise to stable oscillations in density perturbations. We can think of these oscillations as the pressure force when attempting to bring perturbations to 0, overcompensates, causing a an overdensity to become underdense.

If the pressure gradient is not constant we need to look at the time-averaged pressure gradient to check if an object is collapsing. This gives rise to a new quantity, the *filtering mass* which is essentially a time-averaged Jeans mass. Due to this it is more stable over large time-scales and gives a better idea of how much growth of overdensities has been suppressed. This is a useful quantity that is often used when looking at halo collapse, which we will come back to in 2.1.2.

On scales $\lambda \gg \lambda_J$ the pressure term is negligible and (1.16) becomes:

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\rho_0 p\delta \tag{1.18}$$

In an Einstein-de Sitter universe(ie. pure cold dark matter universe, which should behave similarily to our universe in the matter-dominated era), this equation would have a solution on the form:

$$\delta(t) = At^{2/3} + Bt^{-1} \tag{1.19}$$

Here we have one term which will diminish for large t, and one which increases. As t increases the diminishing term eventually becoming negligible. We are then left with a term which increases as $\propto t^{2/3}$. This means that any density perturbation will diverge given a large enough t. This simple example shows that on large enough scales compared to the Jeans length, overdensities become more dense and voids will become more empty, consistent with expectations. As we assume the universe starts out, it is nearly completely homogeneous. Due to dark matter being collisionless we can neglect the pressure term and dark matter density perturbations increse as $t^{2/3}$ while baryonic matter was kept at equilibrium over very large regions. Then as the universe cools down and becomes less dense, the Jeans length evolves as roughly(assuming ideal gas and matter domination):

$$\lambda_J = c_s \left(\frac{\pi}{G\rho}\right)^{1/2} \sim \left(\frac{T}{\rho}\right)^{1/2} \propto a$$

Such that baryonic matter, starting with smaller scales and then larger scales will fall into overdensities caused by the gravitational collapse of dark matter, giving rise to the large scale structure of galaxy clusters, walls and voids. Note that the temperature used in this proportionality is the local gas temperature, and thus it gives a poor picure of reality once this temperature deviates significantly from the CMB temperature (in which case linear perturbation theory does not hold either). Once the gas temperature decouples from the CMB, we assume it expands adiabatically, in which case the gas temperature evolves as $T \propto a^{-2}$(as we will show in (3.2.4)) such that λ_J will then on large scales evolve as $\propto a^{1/2}$

1.3.3 Viralization

Once a system as a whole becomes gravitationally stable we say that it has *virialized*. To undestand how stable overdensities behave, it is helpful to understand what a stable system means and what we expect of a stable system.

Virial theorem

Assuming an object only affected by gravity has an angular momentum $\mathbf{l} = \mathbf{r} \times \mathbf{p}$, we can write its change in angular momentum as:

$$\mathbf{r} \times \dot{\mathbf{p}} = -m\mathbf{r} \times \nabla\phi \tag{1.20}$$

using $\mathbf{F} = m\mathbf{a} = -\nabla\phi$, where ϕ is a gravitational potential. If we were to take a scalar product of (1.20) we would simply get the time derivative of the kinetic energy on the left-hand side. By summing this over multiple objects we would get the time derivative of the total kinetic energy. By rewriting the right-hand side as a sum of gravitational contributions between these objects we would see that this is in fact the time-derivative of the total potential energy. This implies that the total energy of the cluster is conserved if we do not add any exterior forces. If we take a scalar product of (1.20) with its position r_i we can write the expression as:

$$\sum_i \dot{\mathbf{p}} \cdot \mathbf{r}_i = \sum_{i \neq j} \sum_j \frac{G m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} (\mathbf{r}_i - \mathbf{r}_j) \cdot \mathbf{r}_i \tag{1.21}$$

If we take the scalar product with respect to $\mathbf{r}_j$ we would get a similar expression. We could then rewrite the left-hand sides to get expressions on the form:

$$\sum_i \left[\frac{1}{2} \frac{d^2}{dt^2} (m_i \mathbf{r}_i \cdot \mathbf{r}_i) - (m_i \mathbf{v}_i \cdot \mathbf{v}_i) \right]$$

The first term here contains the moment of inertia $I = m\mathbf{r} \cdot \mathbf{r}$ of the whole system, while the second term is twice the kinetic energy. If we combine the right-hand side of (1.21) with its eqivalent for $\mathbf{r}_j$ we get the potential energy of the system. Combining this, we then get:

$$\frac{1}{2} \frac{d^2 I}{dt^2} = 2E_{\text{kin}} + E_{\text{pot}}$$

Averaging this over a long time and realizing that if the objects are bound within the system dI/dt must be finite and tend towards zero as time goes towards infinity. We thus arrive at the virial theorem:

$$2\langle E_{\text{kin}} \rangle + \langle E_{\text{pot}} \rangle = 0 \tag{1.22}$$

Here the angle brackets correspond to the time-averaged value. This result holds for non- relativistic particles as long as there are no external forces affecting the system which is a reasonable assumption for most galaxies. This tells us that once the gravitationally bound system is stable we have a relation between the motion in the system and its total energy.

Virialization temperature

One quantity we can compute using the virial theorem is the *viralization temperature* T_{vir}. The time-averaged typical velocity of an object in a virialized system is given by

$$v = \left(\frac{GM}{r_{\text{vir}}} \right)^{1/2} \tag{1.23}$$

where M is the total mass and $r_{\rm vir}$ is the radius of the virialized system. This is simply the circular velocity we would expect of a particle orbiting an object with mass M. Assuming the gas in a halo moves isotropically with this velocity and is in thermal equilibrium we could use the equipartition theorem of energy which is a relation between the internal energy U of a molecule, its degrees of freedom f and the thermal energy associated with this. As we assume that our system consists of primarily of hydrogen and free electrons we expect that they only have the 3 translational degrees of freedom. Writing this out for our system we find the virial temperature:

$$\begin{aligned} U &= \frac{f}{2} NkT \\ \left(\frac{1}{2} + 1\right) Mv^2 &= \frac{3}{2} \frac{M}{\mu m_p} kT \\ T_{\rm vir} &= \frac{\mu m_p v^2}{2k} \end{aligned} \tag{1.24}$$

Where in the second step the virial theorem was used. N is the number of molecules and μm_p is the mean molecular weight of a particle, with m_p as the proton mass. This quantity is important to understand how gas cools down as we will see in section 2.1.

Chapter 2

Dark ages and Reionisation

The last surface scattering marks a significant transition in our universe. Before this, photons had a very short mean free path due to Thomson scattering with free electrons, while afterwards the mean free path of photons suddenly became almost infinite due to the recombination of electrons and protons into neutral atoms. The information gathered from the photons emitted at this point is very important for our understanding of the universe. It gives us information about the conditions in the very early universe and much time and effort has been devoted to deciphering the information contained.

What comes afterwards is in stark contrast. The period known as the dark ages is called so due to how little information can be gathered from this epoch from photons in the visual range of the electromagnetic spectrum. While the CMB initially had a temperature from blackbody radiation corresponding to the colour red, this redshifted into the infrared, such that were we to have existed during this epoch, we would not be able to see anything with our human eyes. We also expect the universe to be significantly more homogeneous during this period than later, such that the gravitational distortions of the CMB should be small. Due to the lack of probes of this epoch we therefore know little of its evolution.

This epoch ends when the first stars and galaxies begin to form, and fusion of hydrogen generates photons with frequencies in the visible and UV range. Once galaxies start producing significant amounts of highly energetic photons capable of ionising hydrogen, we expect them to continuously ionise larger regions of space until the universe is ionised.

2.1 The Dark Age of the Universe

2.1.1 The alpha and omega of the dark ages

When did the dark ages start and when does it end? To answer the first question we need to find out when the number of free electrons became significantly small enough for Thomson scattering to have a small enough cross-section. A naive estimate would be when photons have energy less than the ionising energy $\epsilon_0 \approx 13.7$ eV. This is not the case, due to photons following a Maxwell-Boltzmann distribution which has a long tail on the high-energy end

of the distribution. The sheer amount of photons compared to baryons thus prevents this from happening.

We thus, in principle, have to solve the Bolzmann equation to see when the photons no longer are able to prevent recombination. We expect the population of free electrons to be governed mainly by

$$e^- + p \rightleftharpoons H + \gamma. \tag{2.1}$$

We define the free electron fraction as

$$X_e = \frac{n_e}{n_e + n_H},$$

where n_i is the number density of particle i. Note that we can set $n_e = n_p$ due to the universe having zero total charge. Assuming statisical equilibrium we arrive at the Saha equation:

$$\frac{X_e^2}{1 - X_e} = \frac{1}{n_e + n_H} \left(\frac{m_e kT}{2\pi\hbar} \right)^{3/2} e^{-\epsilon_0/kT}$$

Solving this with respect to the average temperature when $X_e = 0.99$(1% has recombined) we get $T \approx 4250$K. As this equation is only true while the reaction is in equilibrium. Beyond equilibrium this needs to be solved numerically(see [32]). The numerical solution, shown in Figure 2.1, tells us that at $X_e = 0.01$, $T \approx 2450$. This points towards recombination happening somewhere in between $z = 1600$ and $z = 900$ (using the relation found in (1.9)), which would lead to the beginning of the dark ages.

The dark age ended once the first stars and galaxies started forming and emitting energetic photons. As it is not meaningful to ask when the first star in the Universe was formed (due to this being dependent on primordial density fluctuations in an infinite universe), we may rather ask: when would an observer randomly placed in the universe first observe a star? This calculation has been done assuming quantum fluctuations following a Gaussian probability distribution and a Press-Schechter model in [36]. There they get a redshift $z \approx 38$, and the star would have formed out of a 6.3σ fluctuation (corresponding to a probability of $\sqrt{2/\pi} \int_{6.3}^{\infty} e^{x^2/2}\, dx \approx 3 \cdot 10^{-10}$). A newer estimate by [37] predicts that the redshift for the first observable star should be around $z = 65.4$.

2.1.2 First star formation

From section 1.3 we saw how an overdensity grows. Within overdensities we expect there to be a significantly larger chance of halos forming. These will collapse on timescales of $t_{\text{coll}} \sim \lambda_J / c_s \approx (G\rho)^{-1/2}$ then virialize. The gas falls in with a velocity comparable to v from (1.23). Once the gas approaches a radius $r \sim \lambda_J$ or equivalently the total mas $M \sim M_J$, the gas will collide and cause shocks, heating it up until the gas reaches the virial temperature.

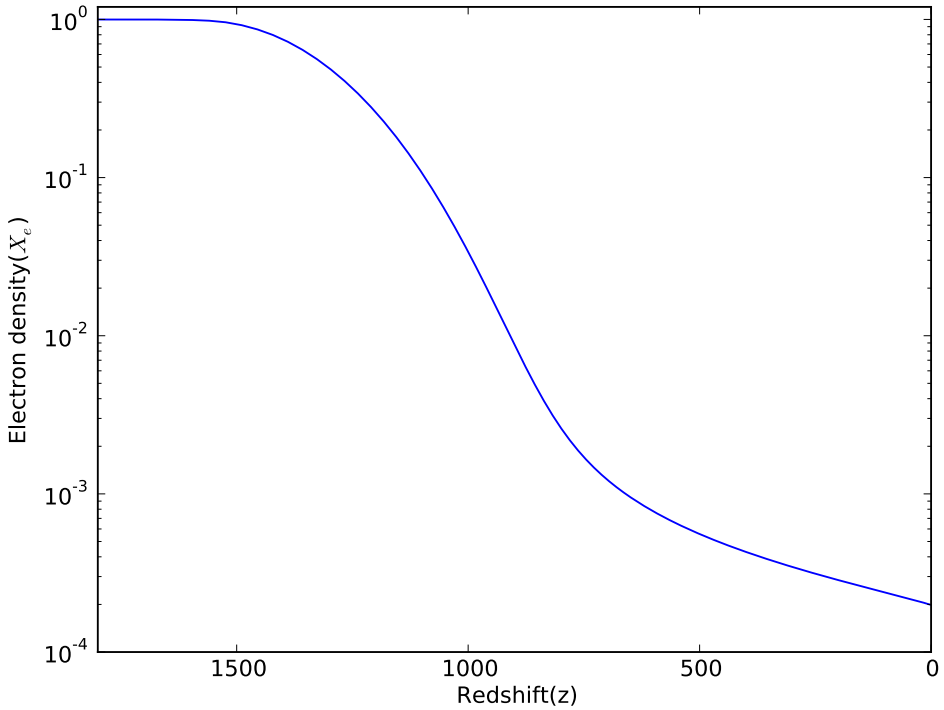

Figure 2.1: The fractional free electron density X_e as a function of redshift z. When the universe was ionised, there were as many free electrons as protons. This number decreases as the energy density of the universe decreases and electrons are captured by baryons, until the photons were decoupled and the dark age of the universe began

Accretion of baryons

We have already mentioned how important it was for the photons to decouple from the baryons, causing cosmologic recombination. Similarly, this coupling affects the baryons and acts as a friction force on any ions moving relative to the CMB frame. This prevents baryons from slowing down enough to fall into dark matter halos until around $z \sim 200$, when the number of electrons is small enough that Compton heating is inefficient. In addtion this very same effect slows down fast electrons, causing a net result of cooling baryonic gas.

To form stars within an overdensity a significant amount of baryons need to fall into the potential well. If we assume a dark matter overdensity has already virialized with the gravitational potential ϕ, we may then ask how much baryonic gas will accumulate. As a simple approximation we can ignore cooling, shocks and Hubble expansion between this overdensity and the surrounding gas (IGM, the *intergalactic medium*). Once the gas is accreted and stable we should expect it to be in hydrostatic equilibrium:

$$\nabla P_b = -\rho_b \nabla \phi \tag{2.2}$$

where P_b is the baryonic pressure. Assuming adiabatic compression of a monoatomic gas,

can write $P\rho^{-5/3} = \text{const}$, which means that:

$$\frac{P_{\mathrm{b},\phi}}{P_{\mathrm{b,IGM}}} = \left(\frac{\rho_{\mathrm{b},\phi}}{\rho_{\mathrm{b,IGM}}}\right)^{5/3},$$

where the subscript ϕ implies gas inside the potential, and IGM is the gas outside. We can then use this to substitute for $P_{\mathrm{b},\phi}$ and integrate (2.2) to get the ratio of densities:

$$\frac{\rho_{\mathrm{b},\phi}}{\rho_{\mathrm{b,IGM}}} = \left(1 - \frac{2}{5}\phi\frac{\rho_{\mathrm{b,IGM}}}{P_{\mathrm{b,IGM}}}\right)^{3/2} \tag{2.3}$$

$$= \left(1 - \frac{2}{5}\phi\frac{\mu m_p}{kT_{\mathrm{IGM}}}\right)^{3/2}, \tag{2.4}$$

where we used the ideal gas law to get the second equality. If we assume for simplicity that the density of the IGM is equal to the average density we can find an expression for the overdensity of baryons in the potential. Using $T_{\mathrm{vir}} = -\frac{1}{3}\frac{\mu m_p \phi}{k}$ we get

$$\delta_b = \frac{\rho_b - \bar{\rho}_b}{\rho_b} = \left(1 - \frac{6}{5}\frac{T_{\mathrm{vir}}}{\bar{T}}\right)^{3/2} - 1 \tag{2.5}$$

Using this it is possible to find an estimate for the minimum mass a dark matter halo needs to have to accrete a specific overdensity of baryonic mass at a given time. Using $\delta_b = 100$ and assuming $\mu = 1.22$ and a top hat model [4] we get:

$$M_{\mathrm{min}} = 5.0 \times 10^3 \left(\frac{\Omega_m h^2}{0.15}\right)^{-1/2} \left(\frac{\Omega_b h^2}{0.022}\right)^{-3/5} \left(\frac{1+z}{10}\right)^{3/2} M_\odot$$

In halos with a baryon content larger than $3{\cdot}10^4\, M_\odot$ gravity will dominate, while for smaller masses pressure will dominate. The halos with baryonic mass below $5 \cdot 10^4\, M_\odot$, however, will be unable to cool down. Therefore, as structure formation starts on small scales, the first objects to collapse will be the ones with masses at approximately $M_b \sim 5 \cdot 10^4\, M_\odot$. The halos with this baryon mass will have a virial temperature $200\,\mathrm{K} < T < 10^3$ K, which will determine the possible ways the halo can cool down, until they become dense enough to form stars.

In this section we have assumed that the dark matter potential is static when compared to the primordial gas. The baryon velocity relative to the dark matter velocity is, however, quite coherent on large scales, such that the gas moves like a wind, so that we cannot neglect the cumulative effect. This can slow down collapse and increase the minimum mass of a dark matter halo to accrete baryons by an order of magnitude, depending on the local baryon velocity.

Hydrogen Cooling

For a galaxy to form, the halo needs to reduce the pressure and allow for higher densities. For this to happen the gas needs a mechanism to eject energy out of the halo, reducing the

temperature, thus reducing the thermal pressure and once again causing gravitational collapse. On large scales this was done through Compton cooling and Hubble expansion, but on smaller scales and with almost completely neutral gas, this is not effective. In our galaxy this is done primarily through collisionally exciting heavier elements and molecules of eg. carbon and oxygen, which have many the rotational and vibrational radiative transitions.

The molecules then de-excite releasing photons which may be able to escape the halo if its frequency is in an optically thin part of the spectrum. Heavier elements(atom number $Z > 2$), however, are formed within stars. In the dark ages these elements were virtually non-existent, and to cool down the halos had to rely on hydrogen. Atomic hydrogen is very inefficient at cooling down for temperatures below $T = 10^4$ K as a lot of energy is required to excite it from the ground state. In contrast, molecular hydrogen, H_2, can be broken(dissociated) by high-energy photons at higher temperatures, but is able to cool gas down to a few hundred degrees K due to its rotational levels(transition between two lowest rotational levels correspond to $T \sim 512\,\mathrm{K}$, and assuming the gas follows a Maxwell-Bolzmann distribution, the high-velocity tail can cool it further). As the first halos had a temperature well below 10^4 K, they had to rely on H_2 cooling.

H_2 in the early universe forms primarily through the following two-step process:

$$\mathrm{H} + \mathrm{e}^- \rightarrow \mathrm{H}^- + \gamma$$
$$\mathrm{H}^- + \mathrm{H} \rightarrow \mathrm{H}_2 + \mathrm{e}^-$$

where the electron acts as a catalyst. As we see in Figure 2.1 the recombination rate(gradient of X_e) is much lower once H^+ and e^- became more scarce. This means that the electrons can be reused to form H_2, and therefore although forming H_2 through this process is rare(and therefore very slow), it is much faster than recombination. Once densities become high enough and recombination becomes efficient again, the free electron fraction is reduced by a few orders of magnitude and thus halos effectively reach a saturation point of H_2. This means that the gas will emit energy at a constant rate. If the timescale of the cooling is greater than the dynamical time of the gas cloud $t_{\mathrm{dyn}} \approx \sqrt{G\rho}$, the halo will be able to stay approximately virialized and will contract with a constant Jeans mass.

This means that $r_{vir} \sim \frac{\lambda_J}{2}$ and the temperature evolves as:

$$T_{\mathrm{vir}} = \frac{\mu m_p}{2k} v^2 = \frac{\mu m_p}{2k} \frac{GM}{r_{\mathrm{vir}}} \sim \frac{\mu m_p}{2k} G M^{2/3} \left(\frac{4\pi}{3}\rho\right)^{1/3} \propto \rho^{1/3},$$

where v is the circular velocity from (1.23). If however the cooling timescale is less than t_{dyn}, the pressure will be to low to prevent gravitational collapse, and the gas cloud will collapse in roughly free fall time $t_{ff} \sim \frac{r}{v}$. For the primordial gas first accreting into a dark matter halo the dynamical time is simply the hubble time $t_H = H^{-1}$, but this changes as the halo becomes more dense.

In high-density gas blobs the H_2 saturation point can be exceeded by generation of H_2 through 3-body processes:

$$H + H + H \rightarrow H_2 + H, \qquad H + H + H_2 \rightarrow H_2 + H_2,$$

increasing the cooling rate, allowing further collapse. When the density approaches the density of the lowest regions of the corona of our Sun ($\sim 10^{12}$ cm^{-3}) H_2 can collide with other atoms or particles, temporarily being able to absorb and emit photons in a continuum (*collision-induced emission*), once again improving the cooling rate until at some point the density becomes large enough to ionise H_2.

As atomic hydrogen requires a higher temperature to cool down the gas, it is only in the more massive halos that this will occur independent of H_2. But as we saw in section 2.1.2 halos of this size will collapse much later. These are however the halos which most probably formed the first sustainable galaxies, as we will see in section 2.2. As a result of feedback we expect the galaxies formed by H_2 cooling and its neighbour galaxies to halt its star production after the very first stars due to H_2 dissociation.

Lyman α radiation, the transition between the ground state and first excited state of neutral hydrogen, may become a more important cooling mechanism in later stages, as the gas may cool if hydrogen is excited by collisional excitation with an electron. Although the mean free path of these photons is very short due to the abundance of neutral hydrogen, the net effect is converting thermal energy into radiation, and we expect a lot of this radiation to be present for $T \gtrsim 10^4$, especially for high densities of free electrons (see eg. [10]).

Properties of the first stars

In the first stars(Pop III) there were no metals. Due to this, they had properties which are significantly different from stars formed more recently(Pop I, third "generation" of stars; characterized by high metalicity). To maintain hydrostatic equilibrium stars emit photons primarily through fusion of hydrogen to helium. If Pop I stars accrete enough gas they start burning hydrogen through what is called the *CNO-cycle*, which due to its efficiency is able to generate a radiation pressure capable of withstanding further gravitational collapse. Smaller stars are unable to reach the temperature required for this, and burn hydrogen through the *pp-chain*.

As the CNO-cycle requires carbon, nitrogen and oxygen as catalysts, Pop III stars are restricted to the slower pp-chain. As a result, more gas is accreted onto the star and Pop III stars tend to be more massive. This in turn means that, like their massive Pop I counterparts, they have a shorter life span. Also, due to lack of radiation in the core (and therefore less radiation pressure, therefore compressing the core), the core temperature is much higher than for a Pop I star. This, combined with that Pop III stars are optically very thick to its own radiation, due to having no metals, cause the effective temperature of the star to be much higher compared to that of a pop I star of similar mass or similar luminosity.

Due to its accretion rate being determined by its radiation pressure, the luminosity will be very close to the Eddington luminosity which is proportional to its mass. The high temperature of the primordial stars means that they are able to produce a lot of high-energy photons. In fact, a large Pop III star will produce $\sim 10^5$ ionising photons per proton, while a typical Pop II star in comparision generates ~ 4000 ionising photons per proton during its lifetime, or $\sim 50\%$ more ionising photons per unit mass. This means

that to ionise the universe completely you would in theory only need to have 10^{-5} of the protons in the universe within stars, if these stars were all Pop III.

Fragmentation

We have so far explained how as a halo accretes gas, it needs to cool down to achieve a higher gas density, until at some point stars may form in the densest regions. We have however not explained how multiple stars, instead of just one massive star, form from a single gas clump, as one would naively expect. This is what is called *fragmentation*. If one takes angular momentum into account we would expect the gas to form a disk, as gravitational collapse typically collapses at different rates. If the disk is compressed by gravitational collapse, conservation of angular momentum requires that the rotational velocity increases. This will in turn prevent further collapse.

We therefore have a two-dimensional equivalent of the Jeans length preventing collapse of the disk if r is too small, and a critical rotational velocity length preventing collapse if the initial r is too large(ie. a large angular momentum). This means that if there is some radius of the disk where $r_{\lambda_J,\mathrm{2D}} < r < r_{\mathrm{rot}}$, this would be unstable, and the area $r \geq r_{\mathrm{rot}}$ would continue to accumulate mass which cannot collapse further due to too high angular momentum, causing fragmentation in the disk as done in a simulation by [9]. Fragmentation leads to stars having to "compete" for gas, and tends to prevent too large stars from forming. Fragmentation may also be caused by having different cooling times in different regions causing instabilities, or by turbulence, but this will not be discussed in this document.

2.2 The Reionisation of the Universe

After the first stars form they will start emitting ionising photons, which in turn will ionise it surrounding gas. Eventually as enough stars form the ionised bubbles will become larger and more numerous until at some point the universe is once again ionised. This marks an important stage in the universe for several reasons. Free electrons once again become abundant, and the main radiative transitions of the hydrogen, the Lyman series, have a much longer mean free path. Baryonic matter, which up until now, was negligible in the larger cosmologic picture compared to dark matter, now starts to affect larger structures. This is arguably the last great phase transformation of the universe where cold neutral gas on the order of a few degrees kelvin, is transformed into hot ionised gas at approximately 10^4 K.

As with the dark ages, its duration and end is not clear. In addition to the processes involved being complex and sensitive to physics on many scales the initial conditions from the dark ages are also unknown. Some calculations have predicted that the first observable galaxy due to atomic cooling may have formed around $z \sim 47$ [37], which emphasizes the sensitivity to initial conditions. However, recent observations seem to point towards reionisation lasting for $\Delta z \leq 4.4$ [50] and ending around $z \approx 6$(see eg. [5, 34]), although

there are still some uncertainties related to whether stars or black holes are the primary cause of the ionisation.

2.2.1 Feedback

The process of output of a system being reused as input to the same system, or *feedback*, is critical to understanding the universe during reionisation. Feedback will however continue to play an important role in regulating both galaxy formation far beyond cosmological reionisation, as it will push the Jeans mass upwards by heating up the gas in the IGM, and star formation, through winds and outflows. As we will see, feedback is non-trivial, and the same effect may both speed up and slow down stellar evolution in different parts of the universe. We have already seen how pressure acts as a feedback mechanism to prevent too sudden collapse, and thus slow down the evolution of the halo, and we will in this section emphasize the effect of stars on their surroundings.

H_2 feedback

The first stars were dependent upon H_2 cooling to collapse rapidly. These very same stars, however, as we saw in section 2.1.2 are able to produce a lot of ionising photons. This may dissociate a large fraction of H_2 thus preventing further collapse, and slowing down stellar formation. In fact, as atomic hydrogen has an ionising energy of 1 Ry = 13.6 eV, while H_2 gas can be dissociated at energies down to 11.2 eV(To dissociate H_2 directly one needs 14.7 eV, but if one excites the molecule into a Lyman or Werner line it may decay via the vibrational continuum, causing dissociation, see [46]), it has been shown that the UV radiation from both stars and black holes capable of dissociating H_2 during reionisation is several orders of magnitude lower than what is required to ionise the universe [24]. Adding the fact that the photons within the range 11.2-13.6 eV (*Lyman-Werner photons*) which cannot ionise HI, have a long mean free path, means that neighbouring halos will also have reduced H_2 cooling. The production of stars in the region of a Lyman-Werner source has to wait until halos with $T_{\rm vir} \gtrsim 10^4$ collapse, as the critical virial temperature required for cooling is raised. A HII bubble will start to form around star-forming systems, and any halo within this region will be comprised of only ionised gas. In the systems with a lot of ionised hydrogen, deuterium will start to play a larger role in the cooling process.

In contrast, if a gas cloud is very dense ionising photons may cause an increased H_2 formation rate. This is because as the number density of HI relative to H_2 $n_{\rm HI}/n_{\rm H_2} \gg 1$. Most ionising photons will ionise HI, increasing the number of electrons, which we in section 2.1.2 saw was the main limiting factor. To maximize this positive feedback effect one needs a lot of ionising photons compared to Lyman-Werner. Due to this it has been proposed that the universe may have been ionised primarily by black holes(we will return to this discussion in section 2.2.2).

In systems with $T_{\rm vir} > 10^4$ we expect there to be a large H_2 formation rate and dissociation rate due to a high free electron density. As the system cools down through HI cooling, however, the dissociation rate decreases significantly and almost independent of initial den-

sity and temperature of the object, the H_2 fraction will be fixed around $x_{H_2} \approx 2 \cdot 10^{-3}$ as long as the resultant temperature remains below 3700 K [39](This number is related to the ratio of time scales of cooling, recombination and dissociation).

Photoevaporation

Another internal negative feedback mechanism is connected to the *ionisation front*, the edge of the ionised region. The behaviour of the ionisation front is computed for an analytical case in [15] where the slope of the density profile determines how large the bubble can become, given a star with a constant ionising photon production rate, before its expansion is determined purely by thermodynamic considerations, and expands with the sound speed.

The ionisation front is where the number of ionisations equals the number of recombinations. The degree of ionisation goes from approximately zero to approximately one over the order of the mean free path of ionising photons. We can use this to compute the radius of the ionised region as function of the ionised photon production rate $\dot{Q}$, referred to as the Strömgren radius R_s (the radius of a Strömgren sphere) :

$$\begin{aligned} \dot{N}_{\text{ion}} &= \frac{4\pi R_s^3}{3}\alpha(T)n_e n_{\text{HII}} \\ R_s &= \left(\frac{3\dot{N}_{\text{ion}}}{4\pi\alpha(T)n_e n_{\text{HII}}}\right)^{1/3} \end{aligned} \tag{2.6}$$

Here α is the recombination coefficient. Along the ionisation front electrons are photoionised, increasing the temperature and pressure proportional to the density. If the density profile is steep we thus will also get a sharp pressure gradient which is able to push gas out of the halo.

This process may empty a halo (and it neighbours) of almost its entire baryonic content, dependent upon the geometry of the halo. If halos are hit by an ionisation front this may also increase star formation as the photoionised electrons may trigger the production of H_2, which can cool the gas and cause collapse or fragmentation.

Other effects, such as the generation of radiation pressure from photons created from recombinations within the ionised bubble may also displace neutral gas through scattering. The photons will typically photoconvert to a Lyman-α photon to which the ionised hydrogen is no longer opaque, so that it will escape the bubble and scatter on neutral gas, pushing the gas away from the halo and IGM.

Winds

Massive stars have a correspondingly shorter lifespan, and as Pop III stars were very massive a lot of them may have exploded as supernovae during reionisation and thus affected the stellar formation in the region. For a star within an ionised region at the

Eddington limit in hydrostatic equilibrium we can write:

$$\frac{GMm_p}{r^2} = \frac{L_E}{4\pi r^2}\frac{\sigma_T}{c} \tag{2.7}$$

$$\begin{aligned} L_E &= \frac{4\pi GMm_p c}{\sigma_T} \approx 6.3M\,\mathrm{W} \\ &= 3.2\cdot 10^4 \left(\frac{M}{M_\odot}\right) L_\odot \end{aligned} \tag{2.8}$$

where σ_T is the Thomson cross-section for scattering of an electron. Comparing with our Sun we can get a first order estimate of the lifetime of a $100\,M_\odot$ star.

$$t_{\text{life}} \approx \frac{M/M_\odot}{L_E/L_\odot} t_{\text{life},\odot} = \frac{t_{\text{life}}}{3.2\cdot 10^4} \approx 0.3\ \text{Myr}$$

If we compare this to the duration of reionisation, $t \approx 0.5\,\text{Gyr}$ (assuming reionisation lasts for $\Delta z = 4.4$ and ends at approximately $z_{\text{end}} = 6.0$), we see that there is ample time for stars of this size to end as supernovae during this period.

When these massive stars end their life as supernovae they inject $\sim 10^{44}$ J into the ISM (*interstellar medium*, medium within galaxies, nut not associated with any particular stellar system). This energy is then either converted into heat, and may then cause photoevaporation, or it is converted into momentum and may directly drive gas away. Like with photoevaporation this will limit the amount of remaining gas in the halo to a few percent, and halt star formation. The remainder is heated up and pushed to lower densities. Galaxies may produce winds with similar effects by converting its radiation pressure into momentum.

On the other hand these supernovae winds will spread metals which were produced in the stellar core out into the halo or neighbouring halos, providing a more efficient cooling mechanism than H_2 or HI cooling at low temperatures, speeding up the process of gravitational collapse, and the birth of Pop II stars. The shock may also cause fragmentation or compress gas clouds triggering more star formation.

Thus a neighbouring halo may lose baryons from low density regions and speed up the collapse of high density regions which are able to self-shield from the ionising radiation. The blastwave of the supernova travels faster than the ionisation front generated, and therefore the transpiring starformation may be free of metals.

One therefore expects bursts of Pop III stars, with longer breaks in between where the gas collapses into the halos again.

2.2.2 Reionisation evolution

The reionisation of the universe is, at least conceptually, quite simple. Stars form, then start ionising until at some point the universe is ionised. The physics involved here is however, complicated and several central questions remain open. How did ionisation evolve, which size distribution of halos drove the reionisation, and how important were black holes

during this epoch? A lot of work has been done in this field, and the next generation of observations is expected to shed some light on these questions.

It is common to assume that ionising photons are produced inside galaxies and therefore depend on the star formation rate. If we, as a first approximation, set an average ionising efficiency of galaxies, implying that the total amount of ionised volume is in some way proportional to the volume where mass has collapsed into virialized objects capable of star formation

$$V_{\text{ion}} = \zeta V_{\text{gal}}, \tag{2.9}$$

and the volume within halos is some fraction of the total volume $V_{\text{gal}} = f_{\text{coll}} V_{\text{tot}}$.

If we assume that H_2 cooling is ineffective during reionisation, we need to require that only objects which have a mass corresponding to at least $T_{\text{vir}} = 10^4$ and therefore are able to cool down using atomic cooling to form stars, can be included in V_{gal}. We also implicitly assume that each halo has its own ionising bubble without overlap, which is typically not the case. This is, however, not too problematic as ionising photons in overlapping regions will nonetheless contribute to the combined ionised bubble.

If we further take into account that each hydrogen might recombine, such that to both ionise larger regions of space and keep the region ionised, the ionised region must be smaller by a factor of $N_{i/H} = (1 + \bar{N}_{\text{rec}})$ compared to when assuming no recombination, where N_{rec} is the mean number of recombinations required to keep an atom ionised. If the number of ionising photons required to hold a region ionised $N_{i/H} n_H V$ equals the number of ionised photons ionising the region $N_{\gamma/H} n_H V$ then the region must be ionised. The condition for complete reionisation therefore simply that the following condition:

$$N_{i/H} = N_{\gamma/H} \tag{2.10}$$

holds globally.

N_{rec} can be expressed as:

$$N_{\text{rec}} = \frac{1}{N_H} \int_0^t dt \int_V \alpha(T) n_e n_{\text{HII}} \, dV \equiv \int_0^t \frac{dt}{\bar{t}_{\text{rec}}} \tag{2.11}$$

This is often parameteraized using a *clumping factor* given as [28]:

$$C_{\text{HII}} = \frac{\langle \alpha(T) n_e n_{\text{HII}} \rangle_V}{\alpha_{\text{ref}} \langle n_e \rangle_V \langle n_{\text{HII}} \rangle_V}$$

Where $\langle \rangle_i$ here indicate an average over i, and α_{ref} is a recombination rate reference value. We can then write the average recombination time as:

$$\bar{t}_{\text{rec}} = \frac{1}{\alpha_{\text{ref}} C_{\text{HII}} \langle n_e \rangle_V \langle x_{HII} \rangle_M}$$

Equivalently we can derive a differential equation for the evolution of reionisation by introducing the fraction of volume ionised, the *filling factor* Q_{HII}. We can divide our expression from (2.9) by the total volume to get the filling factor:

$$Q_{\rm HII} = \frac{V_{\rm ion}}{V_{\rm tot}} = \zeta \frac{f_{\rm coll}}{1 + N_{\rm rec}} \tag{2.12}$$

The proportionality factor ζ may be given by

$$\zeta = \mu f_{\rm esc} f_\star N_{\gamma/b} \tag{2.13}$$

with μ being the mean molecular mass, which for a neutral medium can be written as $\mu = 1/[1 - (3/4)Y]$ where Y is the mass fraction of helium, $f_{\rm esc}$ is the escape fraction of ionising photons from its host galaxy, $f_\star$ is the star formation efficiency, and $N_{\gamma/b}$ is the number of ionising photons per baryon. Both $f_\star$, $f_{\rm esc}$ and $N_{\gamma/b}$ are dependent on the stellar population and halo properties. $f_{\rm esc}$ simplifies the computations significantly as one will not have to look at the messy details of radiation transfer and recombinations on ISM scales around stars and black holes.

Our differential equation for the time evolution of the filling factor, which in principle is equivalent to the evolution of $N_{\gamma/H}$ towards the equality in (2.10):

$$\frac{d}{dt}\left(Q_{\rm HII}\right) = \zeta \frac{df_{\rm coll}}{dt} - \frac{\bar{Q}_{\rm HII}}{t_{\rm rec}} \tag{2.14}$$

The bubble will thus expand if $\zeta(df_{\rm coll}/dt) > \alpha C_{\rm HII} \langle n_e \rangle_V \langle x_{\rm HII} \rangle_M$. Recombinations will set an upper limit to how large a bubble can become corresponding to the mean free path of an ionising photon. This criterion can also be used to evaluate whether ionising sources are able to keep the universe ionised.

To compute (2.14) there are some caveats that need to be taken into consideration. Several of the quantities included in ζ are very model dependent. Halo properties and stellar evolution and populations are by no means well understood at these early times, and ζ should also in principle account for feedback effects. This means that to get a reasonable estimate of ζ can be very computationally expensive.

This also applies for the clumping factor $C_{\rm HII}$ which in principle needs to trace the gas distribution with good enough resolution on low scales, and one needs to be able to compute the geometry of the bubbles, as the average is only over the ionised volume. This means that self-shielded clumps of neutral hydrogen *within* regions of ionised hydrogen should not be included.

The gas distribution of the IGM is primarily located in a cosmic web of gas plus some dense virialized clouds unable to form stars. To get a good estimate of $C_{\rm HII}$ we therefore need to have a very large simulation boxes able to sample large HII regions, in addition to resolving the gas clumping. One also needs to be able to model the effect of clumping on these systems. Given these problems with the clumping factor one has had to selectively choose the most important physics, which is still not entirely clear.

Another complication comes from the recombination rate α. This does not depend only on temperature, but also on whether recombinations are allowed to go directly to the

ground state, which would be expected in the low-density IGM, or if is prevented from this due to the photons then being immediately absorbed again, giving a zero net ionisation rate. This may happen in high density systems with partially neutral gas (called *Lyman-limit systems*). There the photons will only effectively recombine if they are photoconverted into several non-ionising photons.

Reionisation phases

One can divide hydrogen reionisation into roughly 3 phases. In the early phases of reionisation the bubbles are small; they are of the order of the Strömgren radius from (2.6), and each of the bubbles are sustained by individual galaxies. These are primarily in the highest density regions where we expect star formation to start first. We expect the ionisation to be relatively slow due to the high density of hydrogen to be ionised, and the high density of ionised particles along the ionisation front, causing an increased recombination rate as can be seen in (2.11). Once the ionisation front reaches low density regions the bubble will expand rapidly, while passing by some blobs of high-density neutral gas which may remain neutral. During this period the ionisation is very sensitive to the distribution of gas, and will expand very inhomogeneously.

In the next phase the bubbles start overlapping, and neutral blobs within the ionised regions are more exposed. At this point it starts becoming less meaningful to talk about Strömgren spheres around galaxies and one needs to look at more global properties of the bubbles and their location with respect to overdensities. One may then observe that the bubbles have redshift-dependent characteristic scales as can be seen in Figure 2.2. Most of the IGM will be ionised from multiple sources during this phase.

During the last phase the ionised regions have become so large that a large fraction of the photons will be absorbed by Lyman-limit systems within the bubble, such that the photons never reach the outermost parts of the bubble, slowing down the expansion.

Due to the inhomogeneity of the universe on smaller scales, combined with the bias of galaxy formation, these phases will not start at the same time everywhere and will not be isochronal, meaning that the last phase may happen at one place in the universe while the first phase is still occuring somewhere else. Also, as reionisation spreads faster in low-density regions, while most of the ionising sources form in high-density regions, one will expect high density regions to ionise first on large scales (*inside-out reionisation*), but on small scales low density regions will ionise first (*outside-in reionisation*).

Black hole contribution to reionisation

Based on observations of galaxies in the early universe, and how the population of galaxies evolves with redshift, we see that if reionisation is driven only by stars then a significant contribution to the ionising background radiation must come from faint galaxies. This to some extent expected, as PopIII stars first start forming in smaller halos first, as these collapse faster in the early universe, and should therefore play a more important role early on. As we have seen, PopIII stars also produce a greater amount of ionising photons per

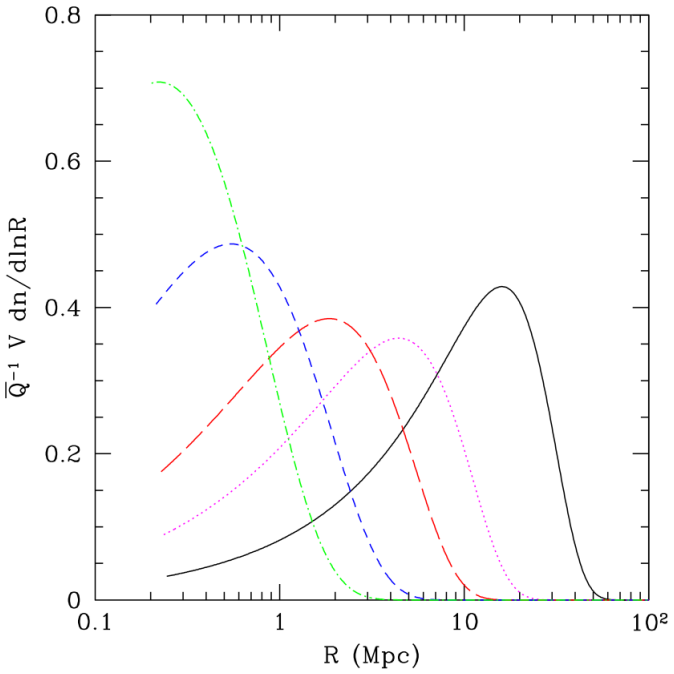

Figure 2.2: Bubble size distribution from [18] at redshifts $z = 18$, 16, 14, 13 and 12 with peaks from left to right which have corresponding filling factors $\bar{Q} = 0.037$, 0.11, 0.3, 0.5 and 0.7. $\zeta = 40$ was used in this model. Each curve has a peak indicating the most common bubble size at the given redshift.

baryon. The question then arises as to how much of the luminosity needed to generate the ionising background is caused by stars?

When the WMAP results came out the results seemed to point towards a very early reionisation, which could not be explained with only stars. To explain this one needed sources that would heat the universe up quickly. The suggested solution was mini-quasars, "small" black holes with masses $M_{\rm BH} < 10^6 M_\odot$. These may have very "hard" spectra (many energetic photons, more X-ray photons relative to number of UV-photons), and would thus have an excess of energy after ionising hydrogen to heat up the universe, as opposed to stellar spectra which are effective at ionising, but have much less excess energy to heat the surrounding gas. This has the added benefit of increasing the number of free electrons, thus boosting the population of molecular hydrogen and thus speeding up cooling, gravitational collapse and early star formation. While the ionisation fronts of mini-quasars are similar to those of stars, the heating reaches further into the IGM, and the escape fraction $f_{\rm esc}$ is assumed to be quite large compared to its stellar counterpart.

Although still unresolved, consensus seems to point towards mini-quasars having a more peripheral role in reionisation, partially as not too many have been discovered at high redshifts. By extrapolating the observations we have of early galaxies to lower luminosities, even assuming very faint galaxies are suppressed due to inefficient cooling or winds, results are compatible with a pure stellar driven reionisation, although the uncertanties still are

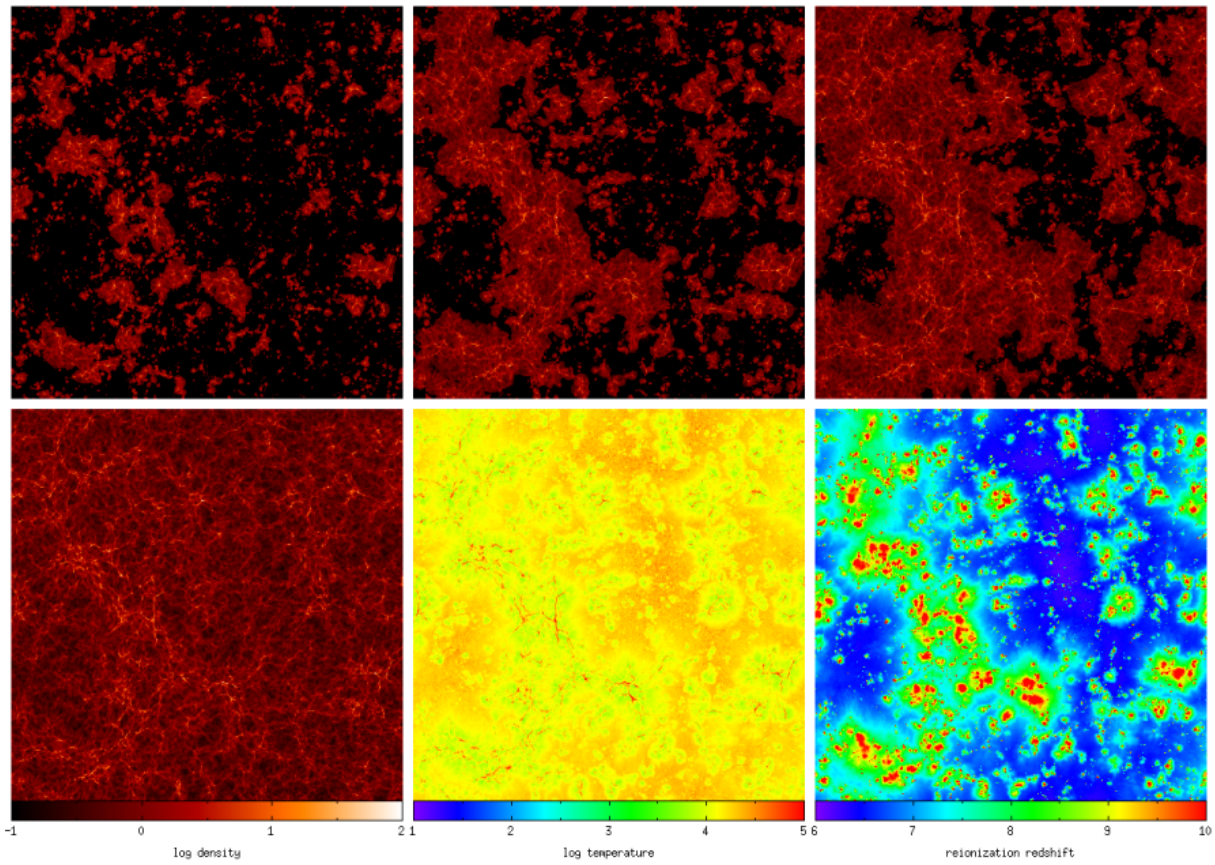

Figure 2.3: Snapshots from simulations from [49] of a slice with width 140 comoving Mpc. The top three and bottom left show the simulation when $\bar{Q} = 0.25$, 0.50, 0.75 and 1.00 of the simulation volume is ionised respectively. Galaxies are formed primarily in the overdensities where their ionisation bubbles merge into larger bubbles. The bottom middle image shows the temperature at the end of reionisation, and the bottom right shows when the regions became reionised. Dense regions and regions close to these are reionised first, then the regions further away. At the end of reionisation the regions which were ionised last are generally hotter as they have not been able to cool down through Hubble expansion.

large. For mini-quasars to be the primary source we would need to assume very hard spectra, and most of the luminosity would have to come from very faint objects. This is discussed in further detail in eg. [5, 51].

2.3 Observational probes

When observing the universe at lower redshifts, we can observe the Lyman transitions of neutral hydrogen. Along the line of sight light from a quasar or other strongly emitting objects the photons will travel through the IGM and more massive objects. This will cause absorption features to appear in the spectrum if the gas emits less photons in that frequency and direction than it absorbs. As Lyman-α (Lyα) is by far the most prominent

line due to the amount of neutral hydrogen in the universe, we therefore expect to see drops in the luminosity where photons redshift into Lyα as it traverses through a region with a high density of HI. This gives rise the what we call the *Lyα forest*.

When looking at spectra from redshift $z \gtrsim 6$, however, we observe what is called the *Gunn-Peterson trough*; light with wavelength shorter than Lyα is missing, as it has been absorbed by neutral gas. This can be used to estimate the end of reionisation, as this shift from the Gunn-Peterson trough to the Lyα forest is due to the IGM ionising, thus limiting the possible regions where Lyα resonance and HI gas coincide.

From observations of Lyα emitters (LAE,observed galaxies, selected due to high Lyα emission) we see a reduction in luminosity, which is not due to a reduction in the luminosity function, for $z \sim 6$. This indicates that there are not significantly fewer LAE's at higher redshift, but that the incoming flux from these are much lower. A similar trend is seen with Lyman break galaxies(LBG, observed galaxies, selected due to significant drop in emission around Lyman limit).

Another important constraint is of the duration of reionisation by the CMB. The coupling of the CMB to baryons is sensitive to the amount of ionised hydrogen/free electrons. With a high density of ions the probability that a photon scatters increases, decreasing the mean free path. The total effect this has on the CMB photons can then be measured.

In principle we should also be able to constrain reionisation by also looking at the first star formation by looking at eg. emission lines of PopIII stars, but there are a lot of uncertainties and very few observations.

Overall, we do have a good idea about the global evolution of reionisation, although there are unresolved problems. Constraints are however improving as we get more data. Some recent constraints are listed in [5], illustrated in Figure 2.4.

When it comes to the dark ages, on the other hand, very little is known. One possible probe would be to look at the CMB anisotropies on the CMB caused by recombination of Lithium atoms at $z \sim 400$. This is however a secondary anisotropy, which will also be difficult to observe due to the redshifted frequency of these photons having a significant foreground (this is discussed in detail in [30]).

A more prominent probe of the dark ages is the 21cm hyperfine transition of the hydrogen atom which we will discuss in chapter 3.

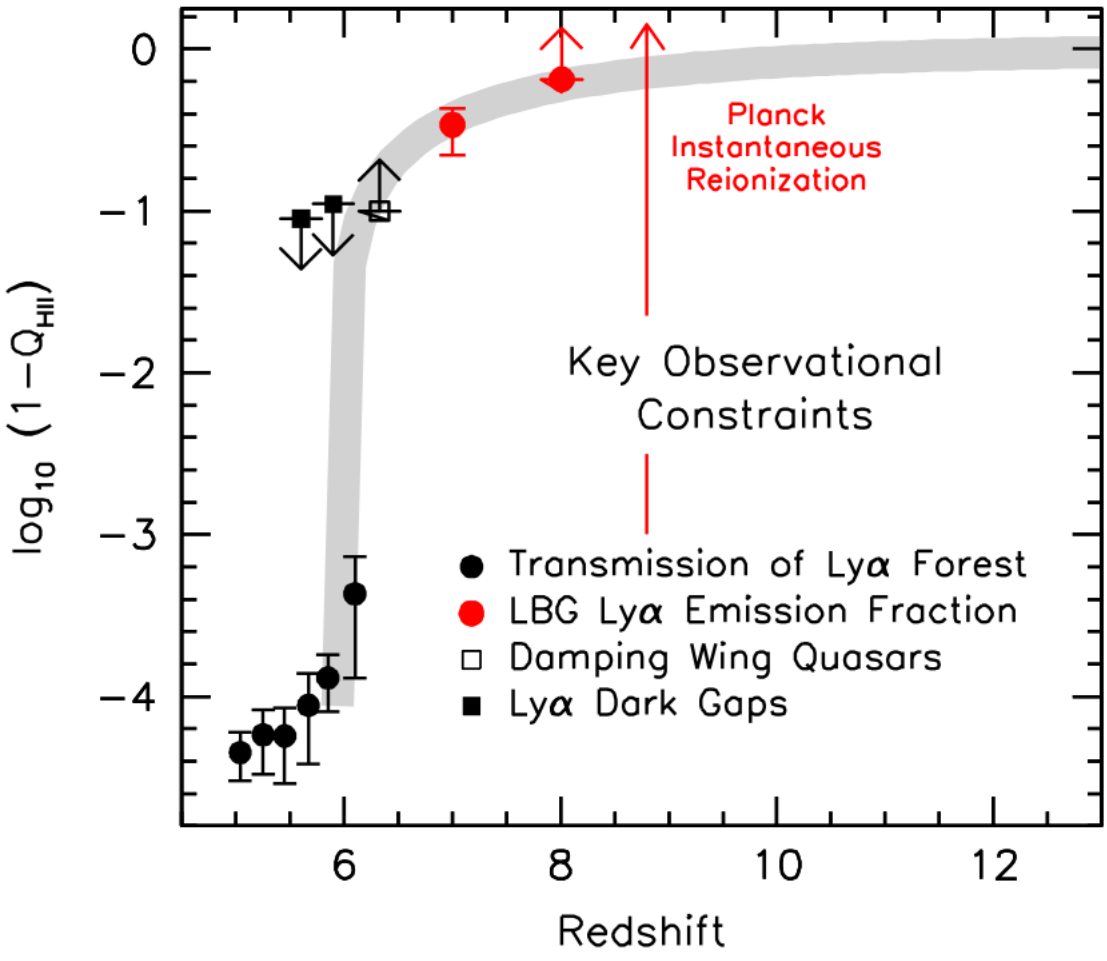

Figure 2.4: ionisation constraints from [5]. $Q_{\rm HII}$ is the ionised hydrogen filling factor. Based on constraints shown here, reionisation is expected to end between $z = 6.5$ and $z = 5.9$.

Chapter 3

The 21cm transition: excitation mechanisms and its cosmological use

Hydrogen is by far the most abundant chemical element in the universe with a total mass fraction of ~ 0.75, and understanding how it behaves is essential to understanding the impact baryons has had on the universe. In fact, for a large fraction of astrophysics, it has until recently been almost the only chemical element which needs to be considered to get a sufficiently accurate understanding of the universe, with minor perturbations due to helium. The hydrogen atom is well understood in quantum physics using the Schrödinger equation. We are able to observe hydrogen lines from objects on a wide range of distances.

Considering this, one might therefore expect that probing the early universe is just taking this one step further. And in a way this is true, but as the universe was neutral before reionisation, there is a similar problem to what happened before cosmological recombination: the universe is optically thick in this frequency range, as we saw in section 2.3 with the Gunn-Peterson trough. To circumvent this problem we look at the *21cm line* of hydrogen, a forbidden transition which we can observe anyway, due to the sheer abundance of hydrogen.

3.1 Atomic levels of hydrogen

3.1.1 Hydrogen states

The hydrogen atom consists of one proton in the nucleus and one electron bound to the nucleus. To model the electron completely we need five quantum numbers:

- The principial quantum number n describes the energy level of an atom. It can have any positive non-zero value. The $n = 1$ state is the ground state, and any larger value implies an excited state. For a time-independent potential like the Coloumb potential of a stationary proton, n also describes the eigenvalue of the Hamiltonian, ie. the energy.

- The azimuthal quantum number l is related to the orbital angular momentum of the system, and gives the subshell of the atom. It can have values of $l \in [0, n-1]$. The subshells are often referred to as "orbital x" where $x = s, p, d, f$ corresponds to $l = 1, 2, 3, 4$. $2(2l+1)$ electrons can fit into one subshell. Magnitude of Angular momentum is given by $L = \hbar\sqrt{l(l+1)}$.

- The magnetic quantum number m_l refers to the projection of the angular momentum typically along the z-axis $L_z = m_l\hbar$ and can have values $m_l \in [-l, l]$.

- The spin quantum number s is an inherent property of a particle. Fermions like the electron, proton or neutron have spin $s = 1/2$. Magnitude of spin for fermions is therefore $S = \hbar\sqrt{s(s+1)} = \sqrt{3}\hbar/2$.

- m_s is to s what m_l is to l; the projection of spin along the z-axis $S_z = m_s\hbar$, and it can therefore have the values of $m_s = \pm 1/2$.

If we solve the radial Schrödinger equation with a Coloumb potential:

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[-\frac{e^2}{r} + \frac{h^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu, \tag{3.1}$$

we find that the allowed energy levels are

$$E_n = -\frac{me^4}{2\hbar^2}\frac{1}{n^2} = -\frac{e^2}{2a_0}\frac{1}{n^2} \equiv \frac{E_1}{n^2}. \tag{3.2}$$

The ground state has the energy $E_1 = -13.6$ eV, assuming a free electron at rest has zero energy. The transitions from one energy level correspond to a photon wavelength of:

$$\lambda = \frac{hc}{|\Delta E_{if}|} \tag{3.3}$$

$$= \left|\frac{hc}{E_1}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)\right|, \tag{3.4}$$

where subscript i and f correspond to the initial and final states. This picture becomes a bit more complicated once you take the fine structure into account, but is sufficient for our purposes here. Note that the absolute value was used here as this formula would give both the minimum energy required for absorption, but also for emission. The hydrogen transitions with $n_i > 1$ and $n_f = 1$ are called the Lyman series, $n_i > 2$ and $n_f = 2$ are part of the Balmer series, $n_i > 3$ and $n_f = 3$ are in the Paschen series. Of these the Lyman series is by far the most important in astrophysical processes, and also the most energetic with the first least energetic transition in the series being $Ly\alpha$, having an energy of ≈ 10.2 eV which is four times stronger than the strongest Balmer transition.

Einstein coefficients

Absorption is the process wherein a photon is absorbed by an atom, thereby exciting the system to a higher energy state. *Emission* is the process of emitting a photon, deexciting the system. This can be divided into two mechanisms: spontaneous emission in which the emission occurs when the atom interacts with the electromagnetic field, resulting in the excited state not being a true eigenstate, such that the eigenstate is a superpostion of the electron states and the electromagnetic field states, which may give rise to a photon; and stimulated emission, in which one may direct a photon with the wavelength corresponding to the energy difference between the electron state and a lower state towards the system, thus causing the system to release a photon with the same frequency and direction as the other photon.

In the following equations we will use notation common when looking at an ensemble of particles, which is more useful for radiative transfer, but may need some explanation. We will also refer to *excitation* and *deexcitation* instead of absorption or emission to emphasize that these are effects related to so-called *bound-bound* transitions (no recombination or ionisation).

The *transition rate* between two energy states for (incoherent and unpolarized) stimulated deexcitation is given by:

$$R_{i\to f} = \frac{2\pi}{\hbar}|M_{if}|^2 \bar{J}^{\chi}_{\nu_0} \tag{3.5}$$

where p is the matrix element of the electric dipole moment between two states We will go into further detail regarding p_{ab} in Appendix B. $\bar{J}^{\chi}_{\nu_0}$ is the frequency-averaged, angle-averaged intensity:

$$\bar{J}^{\chi}_{\nu_0} \equiv \frac{\int_0^\infty \frac{1}{4\pi}\int_0^{4\pi} I_\nu \, d\Omega \, \chi(\nu-\nu_0)\, d\nu}{\int_0^\infty \chi(\nu-\nu_0)\, d\nu} \tag{3.6}$$

where

$$\int_0^\infty \chi(\nu-\nu_0)\, d\nu = 1 \tag{3.7}$$

is the normalized profile function for induced deexcitation around the frequency of the transition between the states. This takes into account the spread of the energy required for a given energy level, once you take into account natural, thermal and pressure broadening, and the bulk motion.

We may now for simplicity look at a two-state system with l being the ground state with population n_l, and u being the excited state with population n_u. We could then follow the evolution of the population n_u with the following equation:

$$\frac{\partial n_u}{\partial t} = -n_u A_{ul} - n_u B_{ul}\bar{J}^{\chi}_{\nu_0} + n_l B_{lu}\bar{J}^{S}_{\nu_0}. \tag{3.8}$$

The A and B's, called *Einstein coefficients*, correspond to, in the order they appear: the transition probability per second per particle for spontaneous deexcitation, and relations stimulated deexcitation and excitation such that $B_{if}J_\nu^\alpha$ gives the transition probability per second per particle in state i given frequency ν and profile α. Here we introduce the normalized profile function for excitation ς which may be defined as χ in (3.7), with a corresponding $\bar{J}_{\nu_0}^\varsigma$ defined in the same way as in (3.6), while for spontaneous deexcitation we may write the profile function as:

$$\psi(\nu-\nu_0)=\frac{1}{4\pi}\frac{\Gamma}{(\nu-\nu_0)^2+(\Gamma/4\pi)^2}. \tag{3.9}$$

also normalized as in (3.7). Note that this applies to a single atom. For an ensemble of atoms we have to integrate over all velocities to get the Voigt function which is derived in subsection 3.2.3. Γ is the line width associated with the lifetime of the state, which we also will clarify in subsection 3.2.3. We will reserve the use of ϕ as a general transition profile, not associated with a specific type of transition.

Thus everything (3.8) states is that the change in the excited state is equal to how many electrons are excited to this state minus the particles which deexcite from this state.

We now want to know the magnitude of the Einstein coefficients. We know that $R_{i\to f}$ from (3.5) must be equal to $B_{ul}\bar{J}_{\nu_0}^\chi$ and want to find an expression for the other coefficients. The Boltzmann law from statistical mechanics states that if we have thermal equilibrium the ratio between two energy levels is given as:

$$\frac{n_a}{n_b}=\frac{g_a}{g_b}e^{-\Delta E_{ab}/kT} \tag{3.10}$$

Where g_i is the degeneracy of the state i. In addition we know that the intensity must equal the Planck function B_ν:

$$B_\nu=\frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1}. \tag{3.11}$$

If we also take into account that $dn_u/dt=0$ and $\phi=\chi=\psi=\varsigma$ (complete redistribution), where ϕ is a normalized profile function of an unspecified bound-bound process in equilibrium we get the Einstein relations:

$$\frac{B_{ul}}{B_{lu}}=\frac{g_l}{g_u} \tag{3.12}$$

$$\frac{A_{ul}}{B_{ul}}=\frac{2h\nu^3}{c^2} \tag{3.13}$$

which must also hold outside equilibrium as they are independent of temperature. The magnitude spontaneous deexcitation transition probability A_{ul} is independent of pressure and radiation field as well, and is a purely atomic/molecular parameter. Typical permitted transitions have $A_{ul}\sim 10^4-10^8$ s^{-1} while transitions with $A_{ul}=1-10^2$ s^{-1} are considered forbidden, and are disfavoured transitions.

We may also define an *oscillator strength* f_{if} such that:

$$B_{if} = \frac{4\pi^2 e^2}{h\nu_{if} mc} f_{if} \tag{3.14}$$

or equivalently:

$$\int \sigma_{if}\, d\nu = \sigma_{if,\nu_0} \int \phi_\nu\, d\nu = \frac{\pi e^2}{mc} f_{if} \tag{3.15}$$

This is a quantum mechanical correction to the classical value of the transition rates we get by assuming the transitions behave like a classical oscillator, and is a good indication of the strength of a given line. In Appendix B we will go into more detail about computing this quantity.

3.1.2 Hyperfine transitions

For an electron we find that the spin interacts with the magnetic field generated by the movement around the atomic nucleus. This causes a split in the energy levels, and this correction, combined with relativistic corrections and quantum oscillations are together referred to as the *fine structure* of the atom. The average energy emitted by an ensemble of hydrogen atoms, however, is still given by (3.2). We use the total angular momentum $J = S + L$ to describe this effect.

In the same way, the nucleus may also affect the energy levels of the electron. This is referred to as the *hyperfine structure* of the atom, and as the name implies, this is an even smaller correction (on the order of m_e/m_p smaller). This is comprised of two effects. The first is related to atoms with the same nucleus charge, but different nucleus number, like the isotopes protium ^{1}H (which is what is referred to as hydrogen in this document) and deuterium ^{2}H. This split however is not within one atom, but between isotopes.

The second effect, which is the one we will be referring to when talking about hyperfine lines, is caused by the spin of the nucleus. The magnetic moment of the nucleus interacts with the magnetic moment of the electron and causes splits in the fine structure. We use define a nuclear spin I, completely analogous to S, as protons also have spin 1/2, and a total angular momentum for the atom $F = J + I$. In the ground state of hydrogen we get a triplet state where the total spin is equal to 1, and a singlet state where the total spin equals 0. This energy difference between the singlet and triplet state can be written as:

$$\Delta E = \frac{4 g_p \hbar^4}{3 m_p m_e^2 c^2 a^4}$$

where $g_p \approx 5.59$ is the g-factor of the proton. If we insert this into (3.3) we get the wavelength $\lambda \approx 21$ cm, corresponding to the frequency $\nu = 1.420$ GHz. This is the 21 cm spin-flip transition. This has a transition rate of $A_{21\text{cm}} = 2.85 \cdot 10^{-15}$ s^{-1}, and is therefore considered highly forbidden and has a lifetime of

$$\tau_{21\text{cm}} = \frac{1}{A_{21\text{cm}}} \approx 11 \text{ Myr},$$

but due to the abundance of hydrogen we are able to observe it. Figure 3.1 shows the splitting of the first two energy levels n. The notation for a given state is $n\,{}_F L_J$.

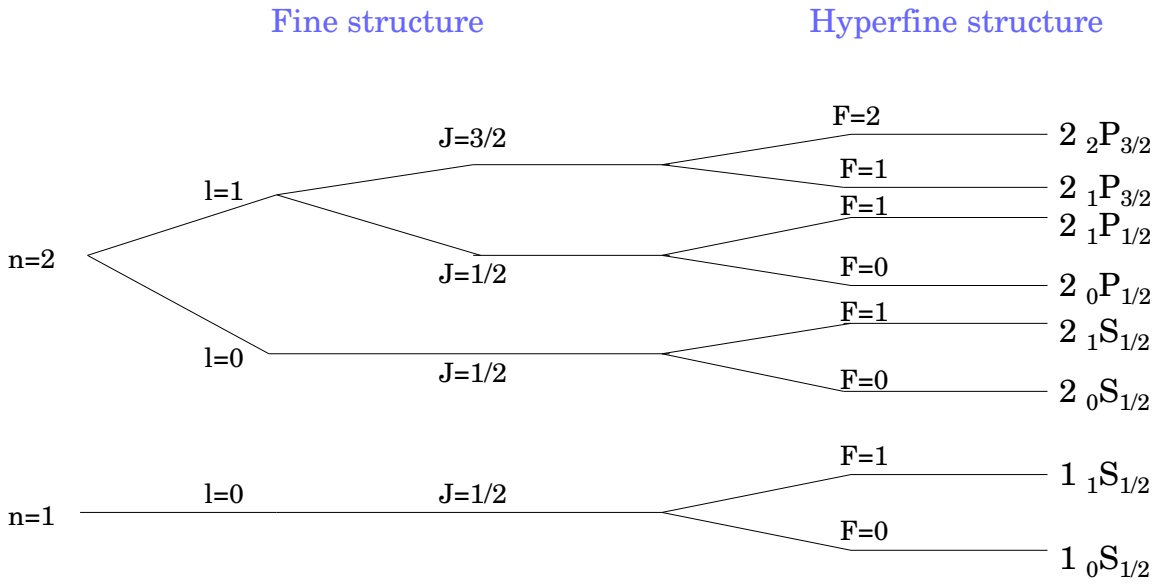

Figure 3.1: The fine and hyperfine splitting of the hydrogen electron energy states $n = 1$ and $n = 2$. On the right we have written the configuration labels using the notation $n\,{}_F L_J$.

3.2 21cm line pumping

We would like to know the population of the hyperfine triplet state of the electronic ground state to get some idea about what we can tell about a 21cm source, based on the intensity of incoming radiation. In this section we will label some hyperfine states as shown in Figure 3.2: 0 for the singlet state, and 1 for the triplet state. Using (3.10) with $g_0 = 1$ and $g_1 = 3$ we may write:

$$\frac{n_1}{n_0} = 3e^{-T_\star/T_S}, \tag{3.16}$$

where $T_\star \equiv \Delta E_{10}/k = 0.068$ K. The spin temperature T_S thus quantifies the ratio of these two populations. As $T_\star \ll T_S$ for most astrophysical applications, almost three out of every four hydrogen atoms with electrons in the ground state are in the triplet state.

In the early universe the spin temperature was closely coupled to the CMB temperature $T_{\rm CMB}$. One of the mechanisms that may drive the population ratio away from $T_{\rm CMB}$ is the *Wouthuysen-Field mechanism.* ,which is central to this work. In addition, collisions between hydrogen and other species may cause an excitation or deexcitation of the hyperfine transitions.

Collisional coupling is most effective in high density regions and at high gas tempeararures T_K. We may define a collision (de)excitation rate rate C_{if} given by the number density n_s of the species s the atomic hydrogen collides into:

$$C^s_{ab} = \kappa^s_{ab} n_s \tag{3.17}$$

$$\frac{C_{ab}}{C_{ba}} = 3e^{-T_\star/T_K} \tag{3.18}$$

With κ as the (de)excitation rate coefficient between the states a and b given by the quantum mechanics involved in the transition. During the dark ages the ionised fraction was very small, and therefore despite the cross section of H–H collisions being very small, this was the dominating effect. In addition H–e and H–H^+ collisions, which have larger (de)excitation rate coefficients (in particular $\kappa^{eH} \gg \kappa^{HH}$), also contribute. Once reionisation starts and the gas temperature increases, e–H collisions becomes more dominant. However, once the temperature starts rising electrons are able to excite the hydrogen to the $2P$ state, generating Lyα, which through the Wouthuysen-Field effect will dominate the spin temperature evolution, mainly due to only $\sim 10\%$ of all collisions leading to hyperfine deexcitations, while each single Lyα photon may scatter many times, before redshifting out of resonance.

3.2.1 Wouthuysen-Field mechanism

In Figure 3.2 we have the 1S and 2P states of hydrogen. Conservation of angular momentum requires $|\Delta l| = 1$ in a transition, while the electric dipole selection rules require $|\Delta F| \leq 1$ except for $F_i = 0 \rightarrow F_f = 0$. An electron in the singlet state $1_0S_{1/2}$ may therefore only excite to the $2_1P_{1/2}$(**3**) or $2_1P_{3/2}$(**4**) states (the blue numbers correspond to the labels used in Figure 3.2). From there it may decay through the emission of $Ly\alpha$ to either of the ground states. Through this process the population of the two states may be mixed, here increasing the population of the triplet state $1_1S_{1/2}$. The process of adding more electrons to the excited state is referred to as *21cm pumping* and its net effect is similar to that of the population inversion done in a laser (which for microwave radiation like the 21cm line would be called a *maser*) , and this increases the chance of spontaneous emission. The Wouthuysen-Field mechanism allows for the opposite to happen as well, depending on the spin temperature. Electrons may be excited to a $2P$ state then down to the ground state. We would therefore like to model the spin temperature, as in some circumstances this may significantly change the expected emission of 21 cm radiation.

3.2.2 The transport equation

The transport equation is used to model how the intensity I_ν at the frequency ν evolves through an optical medium. It can be written as:

$$\frac{dI_\nu}{ds} = j_\nu - \alpha_\nu I_\nu \tag{3.19}$$

where ds is a proper path length element, j_ν and α_ν are the monochromatic emission and extinction coefficents respectively, such that $j_\nu\, ds$ indicates the local contribution to the beam from local emissions at the given frequency, while $\alpha_\nu I_\nu\, ds$ is the amount of photons removed from I_ν through scattering and photon conversion (a process where one photon is absorbed by a photon, exciting the electron, which then cascades down to lower states emitting more than one photon with lower frequencies than the original) and photon destruction. Note that α_ν may be reexpressed in terms of the cross section σ_ν as $\alpha_\nu = \sigma_n u n$, where n is the relevant number density. We can express α_ν and j_ν in terms of the Einstein coefficients for the 21cm line:

$$\alpha_\nu = \frac{hv}{4\pi}\left[n_0 B_{01}\varsigma(\nu) - n_1 B_{10}\chi(\nu)\right]$$
$$j_\nu = \frac{hv}{4\pi} n_1 A_{10}\psi_\nu$$

Rewriting (3.19) in terms of the monochromatic optical depth $d\tau_\nu = \alpha_\nu\, ds$ which effectively weights the path according to how easily photons of the given frequency ν pass through we get:

$$\frac{dI_\nu}{d\tau_\nu} = \frac{n_1 A_{10}\psi_\nu}{n_0 B_{01}\varsigma(\nu) - n_1 B_{10}\chi(\nu)} \tag{3.20}$$

We need to assume thermal equilibrium to be able to relate the intensity to a corresponding brightness temperature T_b. Using the Rayleigh-Jeans approximation $h\nu/kT \ll 1$ our expression simplifies to:

$$\frac{dT_b}{d\tau_\nu} = T_S - T_b.$$

With T_S as defined in (3.16). Before a given beam passes through our optical medium we assume the brightness temperature is given by the background temperature which is typically from the CMB, with the temperature T_{CMB} from (1.10). After the beam has passed through a given the medium the emergent brightness temperature is changed to T_{em}. Integrating we get:

$$T_{\text{em}} = T_{\text{CMB}} e^{-\tau_{10}} + T_S\left(1 - e^{-\tau_{10}}\right) \tag{3.21}$$

The observable quantity we can measure, however, is the redshifted deviation from the CMB-temperature:

$$\delta T_b = \frac{T_{\text{em}} - T_{\text{CMB}}}{1+z} = \left(\frac{T_S - T_{\text{CMB}}}{1+z}\right)\left(1 - e^{-\tau_{10}}\right) \approx \left(\frac{T_S - T_{\text{CMB}}}{1+z}\right)\tau_{10} \tag{3.22}$$

. Here the optical depth of the is given by:

$$\begin{aligned}\tau_\nu &= \int \alpha_\nu \, ds \\ \tau_{10} &\approx \int \phi(\nu) \frac{h\nu}{4\pi} A_{10} \frac{c^3}{2h\nu^3} \frac{3}{4} x_{\mathrm{HI}} n_{\mathrm{H}} \frac{T_\star}{T_S} \frac{da}{H_0 \sqrt{\Omega_m a^{-1}}} \\ \tau_{10} &= \frac{3}{32\pi} \frac{A_{10} h c^3}{k T_S \nu_{10}^2} \frac{x_{\mathrm{HI}} n_{\mathrm{H}}}{H(z)}\end{aligned} \tag{3.23}$$

Where we assumed a matter-dominated universe, and combined $dl = c/(aH)\, da$ with (1.7) to evaluate the line element. The number density in the excited hyperfine ground state was here approximated as $n_1 \approx \frac{3}{4} n_{\mathrm{HI}} = \frac{3}{4} x_{\mathrm{HI}} n_{\mathrm{H}}$. The number density of atomic hydrogen $x_{\mathrm{HI}} n_{\mathrm{H}}$ can be written in terms of cosmological parameters:

$$x_{\mathrm{HI}} n_{\mathrm{H}} = \rho_{\mathrm{cr},0} \frac{\Omega_b (1+z)^3 (1-Y)}{\mu m_p} \tag{3.24}$$

n_0 was reexpressed in terms of n_1 using (3.16) and assuming $T_\star \ll T_S$. We also approximated the profile as a dirac δ at the line centre such that $\phi \sim 1/\nu_{10}$, implying no broadening of the line. (3.23) only applies in a homogeneous universe where the velocity gradient along the line of sight is $H(z)$. A more general expression taking into account the peculiar motion would be

$$\tau_{10} = \frac{3}{32\pi} \frac{A_{10} h c^3}{k T_S \nu_{10}^2} \frac{x_{\mathrm{HI}} n_{\mathrm{H}}}{(1+z) dv_{||}/dr_{||}},$$

where the subscript $_{||}$ indicates line of sight values.

How do we interpret (3.22)? $\delta T_b < 0$ when $T_S < T_{\mathrm{CMB}}$ implying absorption by the medium, while if the opposite is the case there must be emission. Both these are possible during the early ages. Whether we observe the 21 cm line or not is dependent on $T_S \neq T_{\mathrm{CMB}}$ being true. If $T_S \gg T_{\mathrm{CMB}}$ the line will saturate at $T_{em} = T_S$(we can see this by noticing that there is a $1/T_S$ dependence on τ_{10} meaning that the differential brighness temperature $\delta T_b \propto (T_S - T_{\mathrm{CMB}})/T_S$ limiting the maximum positive value), while there is no limit to how large δT_b can become if $T_S \ll T_{\mathrm{CMB}}$.

Thus, we find that the differential brightness temperature is approximately (from [19]):

$$\delta T_b \approx 9 x_{\mathrm{HI}} (1+\delta) (1+z)^{1/2} \left[1 - \frac{T_{\mathrm{CMB}}}{T_S}\right] \left[\frac{H(z)/(1+z)}{dv_{||}/dr_{||}}\right] \mathrm{mK} \tag{3.25}$$

which is on the order of millikelvins.

3.2.3 Line broadening

We have introduced the profile functions ς, χ and ψ in section 3.1.1 as quantities related to the line transitions. These functions give the distribution, or spread of the transition as a function of frequency ν, or equivalently, energy $E = h\nu$. According to the postulates

of Bohr's atomic model, we do however expect quantized energy levels, meaning that we naively should expect the profile functions to have an infinitely sharp dirac δ_D shape. We will here introduce the main causes for this spread in the expected energy for absorption/emission.

Atoms are constantly in motion, both with respect to photons and with respect to eachother, such that the energy available for the transition is not solely given in terms of the radiative energy from the photon, but also the kinetic energy from the atom. In the frequency of the photon will depend on the rest frame of the atom. As the thermal motion of atoms is random on average, this effect, called *Doppler broadening*, will spread the line according to:

$$\nu = \nu_0 \left(1 + \frac{v_{||}}{c}\right). \tag{3.26}$$

The velocity distribution of atoms is given by the Maxwell-Boltzmann distribution such that the number of atoms with velocity between v and $v + dv$ is proportional to

$$e^{-\frac{mv^2}{2kT}}\, dv.$$

We can express this in terms of the frequency ν by using (3.26). Once normalized we arrive at the profile function:

$$\phi(\nu) = \frac{1}{\sqrt{\pi}\Delta\nu_D} e^{-(\nu-\nu_0)^2/\Delta\nu_D^2}, \tag{3.27}$$

assuming the Doppler width $\Delta\nu_D \ll \nu_0$, and is defined as:

$$\Delta\nu_D = \frac{\nu_0}{c}\sqrt{\frac{2kT}{m}}. \tag{3.28}$$

An additional contribution from Gaussian turbulent velocities on small scales compared to the mean free path changes this to:

$$\Delta\nu_{D,\text{eff}} = \frac{\nu_0}{c}\sqrt{\frac{2kT}{m} + v_{\text{turb}}^2}.$$

In (3.9) we already defined the natural profile, dependent on γ, such that $\gamma_n = \sum_{n_l<n_u} A_{n_u n_l}$ with n being the principial quantum number. The mean lifetime of a line is given as $\tau_u = \gamma_u^{-1}$. The broadening causing this, called *natural broadening* is due to the intrinsic uncertainty of the energy and lifetime of a state, given by Heisenbergs uncertainty relation $\Delta E \Delta t \gtrsim \hbar$. If the state is stable, ie. a ground state, then this does not apply. Therefore if both the upper state u and the lower state l are excited states then we have to take that into consideration such that $\gamma = \gamma_u + \gamma_l$

If we define ν_{col} as the average frequency of collisions with other particles, setting the importance of *collisional broadening*, we may write this as a contribution to the natural broadening, giving us the more general expression $\Gamma = \gamma + 2\nu_{\text{col}}$. This effect also temporarily

frees up some energy which may be used by the photon. The collisional effect is more important in dense areas of space, and may be neglected in the IGM.

Other effects which may broaden the line include several interactions with the electric field, but these are not important for our purposes as the electromagnetic field is very weak in the IGM.

Both the natural broadening and the collisional broadening have a *Lorenzian shape* $\propto 1/(1+\Delta\nu^2)$, while the Doppler and turbulent broadening have a Gaussian shape $\propto e^{-\Delta\nu^2}$. If we combine these two, and integrate over the velocities we get what is called the *Voigt function*, parameterized by a damping parameter a and a dimensionless frequency offset x:

$$H(a,x) \equiv \frac{a}{\pi}\int_{-\infty}^{\infty}\frac{e^{-y^2}\,dy}{a^2+(x-y)^2}, \quad a \equiv \frac{\Gamma}{4\pi\Delta\nu_D}, \quad x \equiv \frac{\nu-\nu_0}{\Delta\nu_D} \tag{3.29}$$

Near $\nu = \nu_0$ this is approximately Gaussian if a is small while further away from the line centre, in the "line wings", the Voigt function behaves like the Lorenzian. This expression describes the shape of line transitions well.

If we neglect stimulated emission we may write the cross section at line centre as

$$\sigma_{if,\nu_0} = B_{if}\frac{h\nu_0}{4\pi}\phi(\nu_0) = \frac{f_{if}}{\sqrt{\pi}\Delta\nu_D}\frac{\pi e^2}{m_e c},$$

while the general case where we may write the *Voigt profile* as:

$$\phi(\nu) = \frac{H(a,x)}{\sqrt{\pi}\Delta\nu_D}, \tag{3.30}$$

and the cross section becomes:

$$\sigma_{if,\nu} = \sigma_{if,\nu_0}\cdot\phi(\nu). \tag{3.31}$$

3.2.4 Determining the spin temperature T_S

What drives T_S away from the background temperature $T_{\rm CMB}$? If the medium is only in thermal contact with the CMB we would expect $T_S = T_{\rm CMB}$ after a time on the order of $T_\star/(A_{10}T_{\rm CMB})$ which is much smaller than the Hubble expansion time scale. The other two effects which may prevent this are collisions from other particles on the medium and UV photon scattering. The collisional effects are caused by exchange of electrons between atoms or free electrons, which may change the spin. But after the later stages of the dark ages the Wouthuysen-Field effect (included in the UV photon scattering) dominates as generated Lyα will be able to scatter many times before going out of resonance. These three effects compete in determining T_S. We will mainly follow the calculations of [11, 35] in this section. The evolution of the singlet ground state can be written as:

$$\frac{\partial n_0}{\partial t} = -n_0\left(C_{01}+P_{01}+B_{01}I^S_{\rm CMB}\right) + n_1\left(C_{10}+P_{10}+A_{10}+B_{10}I^\chi_{\rm CMB}\right) \tag{3.32}$$

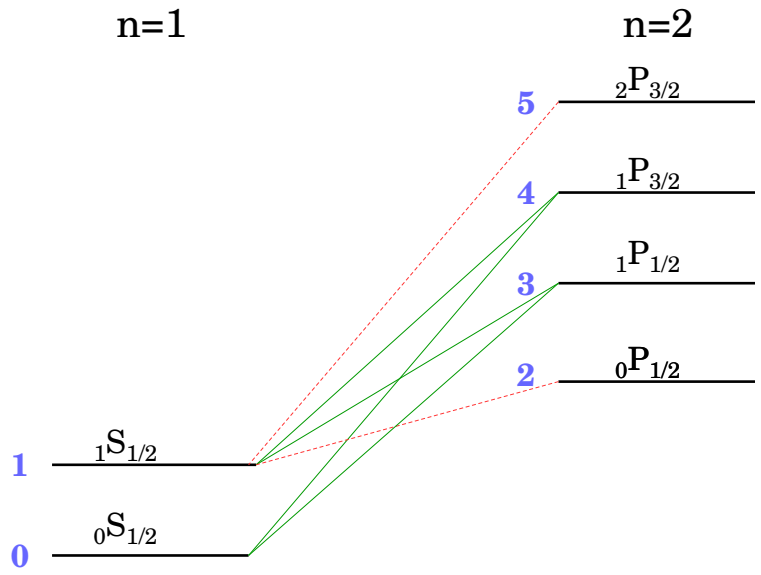

Figure 3.2: abc

Where C_{ab} are collisional (de)excitation rates from (3.17), P_{ab} are scattering (de)excitation rates between states a and b.

In a pure scattering equilibrium we would expect $P_{10}n_1 = P_{01}n_0$. Using this we can define the colour temperature T_c equivalently to how we defined the spin temperature T_S in (3.16):

$$\frac{P_{01}}{P_{10}} = 3e^{-T_\star/T_c} \approx 3\left(1 - \frac{T_\star}{T_c}\right) \tag{3.33}$$

Using the 6-level model described shown in Figure 3.2 we can write the scattering rate between the states (**0**) and (**1**):

$$P_{01} = B_{03}I_{03}\frac{A_{31}}{A_{30}+A_{31}} + B_{04}I_{04}\frac{A_{41}}{A_{40}+A_{41}}$$
$$P_{10} = B_{13}I_{13}\frac{A_{30}}{A_{30}+A_{31}} + B_{14}I_{14}\frac{A_{41}}{A_{40}+A_{41}}$$

We would like to find a way to relate I_{if} and A_{if} to their respective overall Lyα values I_α and A_α. We can relate these by

$$\frac{I_{if}}{I_\alpha} = \frac{g_i}{g_{\text{tot}}}\frac{A_{if}}{A_\alpha}. \tag{3.34}$$

where g_{tot} is the sum of the weights of the upper levels. Using a sum rule for the transitions stating that the sum of all decay intensities from a state nFj to $n'j'$ summing over F and

m_j is proportional to $2j+1$ [1]:

$$\sum_f I_{if} = (2j+1)C$$

Where C is a constant. For states (0)-(5) we have $g_i = 2F+1$ (which we get by counting the possible m_j values of the excited states). To $n'=1, j=1/2$ we then get ($I_{50} = I_{20} = 0$ due to being forbidden transitions):

$$\frac{I_{51}}{I_{41}+I_{40}} = \frac{5}{3}, \quad \frac{I_{40}+I_{41}}{I_{30}+I_{31}} = \frac{3}{3}, \quad \frac{I_{30}+I_{31}}{I_{21}+I_{20}} = \frac{3}{1},$$

while from transitions to $n'=2, j=1/2$ and $n'=2, j=3/2$ we get:

$$\frac{I_{04}}{I_{14}+I_{15}} = \frac{1}{3}, \quad \frac{I_{03}}{I_{12}+I_{13}} = \frac{1}{3}.$$

Combining these equations and (3.34) we get:

$$A_{20} = A_{50} = 0, \quad A_{21} = A_{51} = A_\alpha, \quad A_{30} = A_{41} = \frac{1}{3}A_\alpha, \quad A_{31} = A_{40} = \frac{2}{3}A_\alpha. \tag{3.35}$$

Writing $\chi = \phi$ as the CMB spectrum is well described as a black body and ignoring effects from collisions, as well as absorption and stimulated emission we can write (3.32) as:

$$\frac{\partial n_0}{\partial t} = n_1 A_{10} + \frac{c^2}{2h\nu^3} A_\alpha \frac{2}{9} \left[n_1 \bar{J}_{13} - 3n_0 \bar{J}_{03} + n_1 \bar{J}_{14} - 3n_0 \bar{J}_{04} \right] \tag{3.36}$$

where $\bar{J}_{if} = \int J_\nu \phi_{if}\, d\nu$. We linearize the profile functions assuming that the profiles are identical, but with a frequency offset relative to ϕ_{04} such that

$$\phi_{04} = \phi,\ \phi_{14} = \phi + \Delta\nu_{10}\frac{\partial\phi}{\partial\nu},\ \phi_{03} = \phi + \Delta\nu_{34}\frac{\partial\phi}{\partial\nu},\ \phi_{13} = \phi + \Delta\nu_{10}\frac{\partial\phi}{\partial\nu} + \Delta\nu_{34}\frac{\partial\phi}{\partial\nu}$$

with $\Delta\nu_{ij}$ corresponding to the frequency offset between (i) and (j). We then get (writing $n_1 = n_{HI} - n_0$):

$$\begin{aligned} \frac{\partial n_0}{\partial t} = (n_{HI} - n_0)A_{10} + \frac{c^2}{2h\nu^3} A_\alpha \frac{2}{9} \Bigg[& n_H \int J_\nu \left(2\phi + \frac{\partial\phi}{\partial\nu}\Delta\nu_{34} + 2\frac{\partial\phi}{\partial\nu}\Delta\nu_{10} \right) d\nu \\ & - n_0 \int J_\nu \left(8\phi + 4\frac{\partial\phi}{\partial\nu}\Delta\nu_{34} + 2\frac{\partial\phi}{\partial\nu}\Delta\nu_{10} \right) d\nu \Bigg] \end{aligned} \tag{3.37}$$

(3.38)

which we rewrite to

$$\frac{\partial y}{\partial t} = (1-y)A_{21} + b_1 - b_2 - y(4b_1 - b_2) \tag{3.39}$$

where $y = n_1/n_H$, and b_1 and b_2 are given as:

$$b_1 \equiv \quad A_\alpha \frac{c^2}{h\nu^3}\frac{2}{9}\int J_\nu \left(\phi + \Delta\nu_{34}\frac{\partial\phi}{\partial\nu}\right) d\nu \tag{3.40}$$

$$b_2 \equiv -A_\alpha \frac{c^2}{h\nu^3}\frac{2}{9}\int J_\nu \Delta\nu_{10}\frac{\partial\phi}{\partial\nu}\, d\nu \tag{3.41}$$

At equilibrium we may write:

$$y = \frac{b_1 - b_2 + A_{10}}{4b_1 - b_2 + A_{10}} \tag{3.42}$$

Expressing the Lyα colour temperature in terms of y:

$$\begin{aligned} T_\alpha &= T_\star \left(1 + \frac{1}{3} - \frac{1}{3y}\right)^{-1} = T_\star \frac{3y}{4y-1} \\ &= T_\star \frac{b_1 - b_2 + A_{10}}{-b_2 + A_{10}} \\ &\approx -T_\star \frac{b_1}{b_2} \end{aligned} \tag{3.43}$$

where in the last step we assume $b_1 > b_2 > A_{10}$. We arrived at this assuming collisions and absorption and stimulated emission effects were negligible. Thus *if* Lyα scattering were dominating compared to the other processes we would expect $T_S = T_\alpha \approx T_c$, where T_α is the Lyα colour temperature.

If we go back to the general case from (3.32) in equilibrium we would get an equation similar to (3.39). We could then write y and express it in terms of the spin temperature T_S analogous to (3.43):

$$y = \frac{A_{10} + C_{10} + B_{10}I_{\rm CMB} + b_1 - b_2}{4b_1 - b_2 + C_{01} + B_{01}I_{\rm CMB}} \tag{3.44}$$

$$T_S = T_\star \frac{3y}{4y-1} = T_\star \frac{A_{10} + C_{10} + B_{10}I_{\rm CMB} + b_1 - b_2}{A_{10} + C_{10} - b_2 + (3B_{10} - B_{01})} \tag{3.45}$$

we now want to express b_1 in terms of the total Lyα scattering rate P_α. The Lyα absorption cross section necessary to compute this can be written as

$$\sigma_\alpha = B_\alpha \frac{h\nu_\alpha}{4\pi}\phi_\alpha$$

with B_α being the Einstein B coefficient of the Lyα emission. We may then write:

$$P_\alpha = 4\pi \int \frac{J_\nu}{h\nu}\sigma_\alpha\, d\nu \approx B_\alpha \int J_\nu \phi\, d\nu \tag{3.46}$$

$$= \frac{27}{4}\underbrace{A_\alpha \frac{c^2}{h\nu_\alpha^3}\frac{2}{9}\int J_\nu \phi\, d\nu}_{=b_1} \tag{3.47}$$

writing $B_{01} = 3B_{10}$, $b_2 = -b_1 \frac{T_\star}{T_\alpha}$, and defining a gas temperture T_k analogously to T_S and T_α, we arrive at our final expression for T_S as it was formulated in a paper [13] by George Field in 1958 in terms of the coupling terms y_α and y_k:

$$T_S = \frac{T_\star + T_{\text{CMB}} + y_\alpha T_\alpha + y_k T_k}{1 + y_\alpha + y_k}, \quad \text{where } y_\alpha = \frac{4P_\alpha T_\star}{27A_{10}T_\alpha}, \quad y_k = \frac{C_{10}T_\star}{A_{10}T_k}. \tag{3.48}$$

(3.48) gives us the spin temperature which may be inserted into (3.22), thus making it possible to interpret the results of an observation of δT_b. In the next chapters we will focus on the effect of the total Lyα scattering rate P_α which determines the effect of the Wouthuysen-Field mechanism, but first we will look at the evolution of the spin temperature in the early universe.

Lyman series contribution to T_S

We have so far looked at the effects of the Lyα line on the spin temperature. One would however expect that more energetic photons should be able to contribute to this effect, as the only requirement one would need is that the photons are able to mix the population of the hyperfine states. Although this is true, there are several reasons why we only need to look at the effects of Lyα photons.

Lyα photons scatter on the order of $\sim 10^6$ times before going out of resonance. This is due to the high optical depth around the ground state of hydrogen, causing the mean free path to be much shorter than for other photons. It can be written as:

$$\tau_\alpha = \int \alpha_\nu \, ds = \frac{n_{\text{HI}} c}{H(z)\nu_\alpha} \int \sigma_\alpha \, d\nu = \frac{n_{\text{HI}} c}{H(z)\nu_\alpha} \frac{\pi e^2}{mc} f_\alpha, \tag{3.49}$$

using the expression from (3.15), assuming all neutral hydrogen is in its ground state, which is a reasonable assumption in the IGM. We therefore have $\tau_\alpha \gg 1$ in the early universe, giving a value of approximately the same magnitude as the number of scatterings.

A similar optical depth, however, should also apply to higher Lyn transitions, as they also scatter via the ground state. The main difference is that higher Lyn photons will after few scatterings photoconvert into lower energy photons, such that they will not affect the spin temperature so much directly, but they may produce Lyα photons. If they were to affect the spin temperature directly we would have to have a separate colour temperature T_n equivalently to T_α for Lyα with a strong coupling constant y_n, and the temperature should be driven towards T_k.

Direct Lyn scattering: Following [41] we may show that the Lyn photon contribution is small. The probability that a photon will decay from one state to another can be written as:

$$P_{if} = \frac{A_{if}}{\sum_{f'} A_{if'}}, \tag{3.50}$$

ie. the probability is equal to the transition rate to a given final state, divided by all possible final states. The probability to decay to the $1S$ state from a nP state with $n > 2$ (ie. any Lyman transition except Lyα) is $P_{nP\to 1s} \approx 0.8$ (for more detailed values see Table B.1). In an optically thick medium the photon emitted will quickly excite another electron in the to the same state. This also means that the number of scatterings before it decays to a state which is not the ground state will on average be $N_{\text{scatter}} = 1/(1-P_{nP\to 1s}) \approx 5$. [14] showed that $N_{\text{scatter}} \sim \tau$ if no decay occurs such that we expect

$$P_\alpha/P_n \sim N_{\text{scatter},\alpha}/N_{\text{scatter,n}} \sim 2 \cdot 10^5$$

for any Lyman transition with $n > 2$. This means that the coupling constant must be much weaker than y_α, given that we do not expect a significantly higher amount of any Lyman transition over Lyα.

We can also find an upper limit to how important T_n is in heating T_k by assuming that all momentum energy gained by an atom when a photon is scattered by an atom contributes to increasing T_k. The energy loss of a photon during scatter can be written as:

$$E = \frac{h\Delta\nu_D}{2k_B T}$$

This is a very small number, and the previous argument regarding number of scatterings of Lyn photons compared to Lyα photons can also be used to show that the combined scattering effect is smaller than the heating from Lyα photons. To satisfactorily solve this, one needs to perform a full numerical analysis as done in [6].

Lyn cascade: Lyn photons may decay to the ground state in three different ways. (i) It may emit a Lyn photon directly, which is the case mentioned above. Alternatively it may cascade, emitting more than one photon, including a lower transition Lyn photon. (ii) It may cascade to the $2S$ state, from which it needs to go through a forbidden transition to reach the ground state. (iii) It may cascade to the $2P$ state, from which it will generate a Lyα photon. This final path will increase the number of Lyα photons.

We may ask how important this contribution is. To find out we need to find the probability that a given Lyn photon emits a Lyα photon during its cascade. This is given by the *recycle fraction* f_{recycle} which we may set using the recursive expression from [25,41]:

$$f_{\text{recycle},i} = \sum_f P_{if} f_{\text{recycle,f}}, \tag{3.51}$$

where P_{if} is the decay probability from (3.50). This is simply tracing the cascade backwards finding the total probability that it reaches its final state. To evaluate this expression we first need to set some initial conditions, $f_{\text{recycle,2P}} = 1$ due to this being the Lyα transition, while $f_{\text{recycle,2S}} = 0$ due to the selection rule $|\Delta l| = 1$. Due to the photons emitting in an optically thick medium we need to assume that direct transitions of Lyn will be absorbed again as a Lyn photon, and the net effect of this is zero. Note in particular that

Lyβ the transition $3P \to 1S$ may never form Lyα, as it may only decay to the states $2S$ and $1S$(which would emit another Lyβ photon). For all other transitions we have $0.25 < f_{\text{recycle}} < 0.40$. Ly$\gamma$ may transition to Lyα by cascading to either the $3S$ or $3D$ state, then to $2P$. Some values of f_{recycle} are listed in Table B.1. This means that we may expect that roughly one third of all Lyn photons contribute to the Lyα scattering rate. The remaining two thirds which end up in the $2S$ state will have a very large mean free path once they are emitted through a two photon process, as this is a forbidden transition, and the energies involved are below Lyα. This will however typically amount to less than $\sim 15\%$ of the total scattering rate [25, 41].

Evolution of the spin temperature T_S

To get a good picture of the evolution of T_S we need to understand the relative strengths of the coupling terms and temperatures related to the CMB and the gas. As we already saw from (1.9), the CMB temperature evolves as $T_{\text{CMB}} \propto (1+z)$. Before cosmic recombination the universe was in thermal equilibrium, forcing the spin and gas temperatures to follow the evolution of the CMB temperature through Compton scattering, $T_S, T_k \to T_{\text{CMB}}$. Only once there are too few electrons to couple the gas to the CMB photons at $z \sim 200$ does the gas temperature start to cool adiabatically due to Hubble expansion as $T_k \propto (1+z)^2$ bringing T_k well below T_{CMB}. This proportionality can be found simply by assuming expansion of an ideal monoatomic gas ($\gamma = 5/3$):

$$PV^\gamma = \text{const} \quad \Rightarrow \quad T \propto \frac{1}{V^{\gamma-1}} \propto (1+z)^{3(\gamma-1)}$$

At these early times the high density if neutral hydrogen meant a very high probability of collisions between these atoms, implying $y_c \gg 1$, coupling T_S to T_k. Around $z \approx 70$ density is low enough that $y_c \sim 1$, and T_S starts decoupling from T_k, thus bringing T_S back to T_{CMB}, such that around $z = 40$ we expect $T_S \approx T_{\text{CMB}}$. As the first galaxies form they start heating the surrounding gas, increasing the gas temperature. The ionising sources therefore push the spin temperature upwards, first slowly, then towards the end of reionisation we expect $T_S \to T_\alpha \gg T_{CMB}$. This evolution, however is very sensitive to the ionisation sources. If quasars are important to the ionisation, the reionisation happens earlier, causing the spin temperature to be decoupled from the CMB temperature earlier, thus recoupling T_S to T_k possibly before the gas temperature manages to increase above T_{CMB} as shown in Figure 3.3. Observing the shape of δT_b from this epoch would therefore be an indication of the composition and distribution of ionising sources during reionisation and the dark ages. During late times we expect $T_S = T_k$ as the ionised bubbles become very large and Lyα can scatter on the gas everywhere in the universe.

3.3 Prospects for detection

From our expression for the differential brightness temperature δT_b in (3.25) we see that δT_b should be on the order of ~ 10 mK. We do, however, expect foregrounds from galaxies,

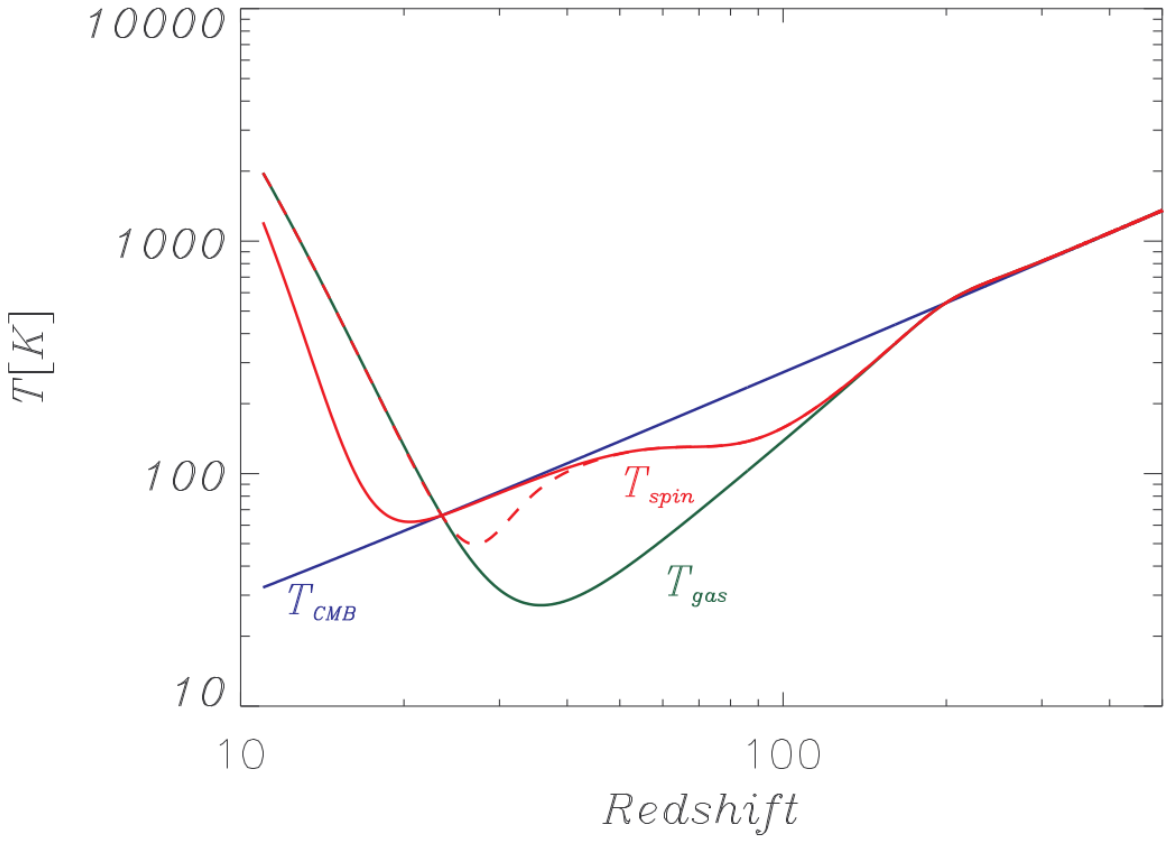

Figure 3.3: The evolution of the CMB temperature $T_{\rm CMB}$, the gas temperature T_k and the spin temperature T_S in the early universe. The solid and dashed lines indicate a model with reionisation dominated by stars or mini-quasars respectively. Figure taken from [51] using simulations from [48]

including our own, from the atmosphere and instrumental noise to make observations very difficult to extract from the data. The frequency range relevant to observing δT_b has until recently been difficult to access due to the physical effects in the ionosphere, and so very little work has been done here. The foregrounds in frequencies $\nu \in [40, 200]$ MHz, which corresponds to redshifts $z \sim 35$–6 respectively, are very high, from some frequencies up to four orders of magnitude higher than the expected δT_b signal. Due to this, earlier 21cm experiments measuring the global/mean signal over the sky as function of frequency have had poor constraints.

Using radio interferometers one may reduce the signal to foreground/noise ratio by 1–2 orders of magnitude. This is due to the large amount of information we can extract from these, which can be used to calibrate the instruments. The downside is the cost involved in conducting these experiments.

Nonetheless, we do expect results from the 21cm observation. As the noise from the different sources have different physical origins, we may model these by independently constraining this. In particular, we may be able to separate the galactic and extragalactic

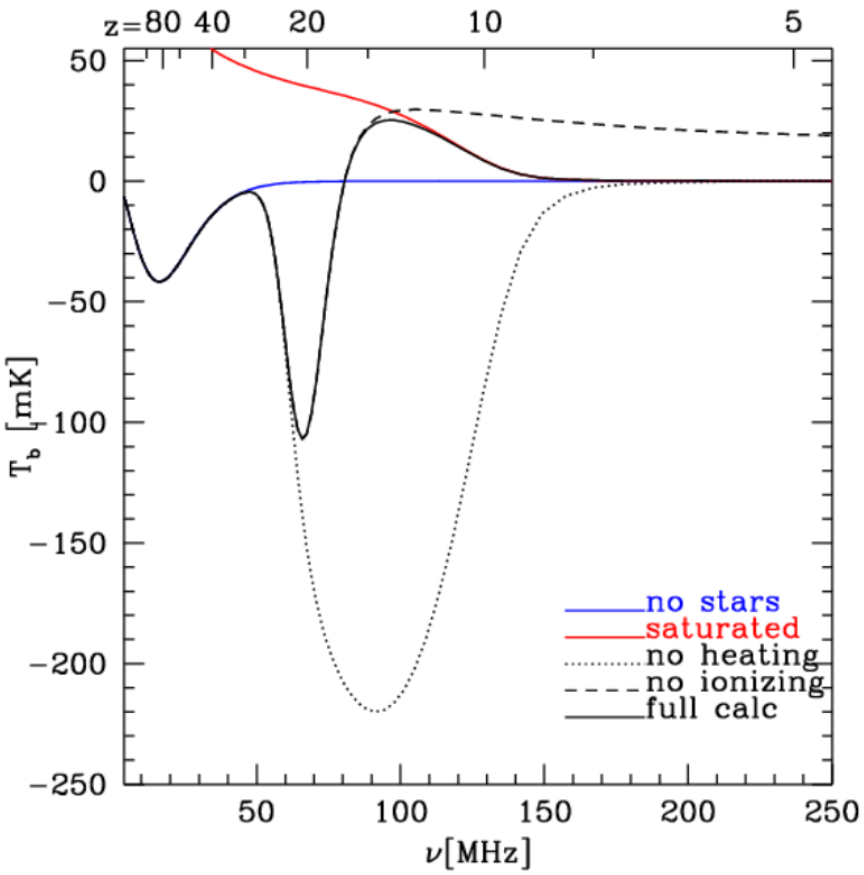

Figure 3.4: The evolution of the differential brightness temperature dT_b for different scenarios from [42].

foregrounds from the signal. We expect the foregrounds to be (spectrally) smooth without any particular features, yet spatially varying. In stark contrast, we expect the 21cm signal to fluctuate spectrally due to the inhomogeneity of the ionised bubble distribution and the structure of the spectrum, yet be spatially smooth.This is however complicated by instrumental thermal noise and calibration issues, and the partially polarized nature of the foreground. To reduce possible sources of error due to polarization and the atmospheric effects, one needs to observe larger region of the sky. One may still end up overfitting or underfitting the foregrounds, resulting in either spurious signals or a significant underestimation of the signal.

LOFAR[1], MWA[2], PAPER[3] and SKA[4] are some of the 21cm interferometry experiments consisting of many smaller antennae, where each individual antenna simply measures the

[1]Low Frequency Array, `http://www.lofar.org/`
[2]Muchinson Widefield Array, `http://www.mwatelescope.org/`
[3]Precision Array to Probe EoR, `http://eor.berkeley.edu/`
[4]Square Kilometer Array, `http://www.skatelescope.org/`

frequency and phase of the incoming radiation. By combining the information from antennae one may get very high angular resolution. By separating the antennae with different distances(baselines), one may observe at different scales.

One goal these interferometric telescopes aim for, is to to extract the *21cm power spectrum*, the variations in intensity of the 21cm line.

To reduce atmospheric noise, DARE[5] will attempt to measure δT_b from space, by orbiting the moon. In addition to avoiding the atmosphere, being on the farside of the moon means one may avoid radio interference from earth and radiation from the sun, meaning that observing the global 21cm signal becomes more feasible, and we can expect to learn more about reionisation and the dark ages.

It has also been proposed to take this one step further by setting up radio telescopes on the moon. This will be necessary to look further back into the dark ages $z \sim 50$–100, as the earths ionosphere is very opaque for $\nu \sim 30$ MHz corresponding to $\lambda \sim 10\,\mu$m due to ozone absorption bands. As the temperature of the moon may reach below 50 K, we may significantly reduce instrumental noise, such that we only need to disentangle the (extra)galactic foregrounds from δT_b.

Once we are able to observe the early universe through the 21cm line, constraints on early star and galaxy formation, ionisation evolution and the matter distribution during the dark ages, elucidating an epoch of the universe we previously have had almost no information about.

[5]The Dark Ages Radio Explorer, `http://lunar.colorado.edu/dare/`

Chapter 4

This thesis

To be able to interpret an observed spin temperature correctly we need to have a good undestanding of how the underlying physics interact. In chapter 3 we saw how T_S is dependent on several components. During the transformation from the dark ages to reionisation, Lyα scattering plays an important role in determining whether neutral hydrogen gas is visible in the 21cm line or not. In this chapter we will therefore look at how we can model the Lyα scattering rate P_α necessary to finding a value of T_S in (3.48).

4.1 Modelling the scattering rate P_α

There is currently no analytical solution to the interaction between a UV source and the spin temperature of the surrounding IGM, but simple models have been made by [41] and [8], whose paper we will focus on in this thesis. In the paper by Chuzhoy and Zheng they assume a spherically symmetric system with a central source emitting UV photons between 10.2 eV and 13.6 eV. Parameters are chosen to be consistent with the IGM around a source with $T_{\text{vir}} \sim 10^4$ at $z \approx 25$. The photons redshift into the Lyman transitions, where the Lyn photons "inject" photons into Lyα resonance. The model thus locates how far a photon travels before it reaches the thermal core of Lyα, where it will continue to scatter within a small region until it redshifts out of resonance. We need not take into account the scatterings before approximate resonance, as the photons reach their asymptotic distribution after few scatterings compared to the $\sim 10^6$ scatterings which take place once they reach the thermal core.

One has to treat the Lyn photons differently from Lyα. During every scattering of a Lyn photon we check whether it photoconverts into a Lyα, while keeping track if its location, irrespective of whether it is close to the line core or not, as all it takes is one scattering to inject a photon into Lyα resonance. For the Lyα we only need to follow its trajectory, but these on average are able to redshift more before reaching resonance, and they scatter more before reaching the thermal core. In the paper these are referred to as "continuum" photons, and their scattering rate is given by:

$$P_\alpha^{\text{cont}}(r) = \tau_{\text{GP}} S_\alpha \int_{\nu_\alpha}^{\nu_\beta} W_\alpha L_\nu \, d\nu \tag{4.1}$$

where τ_{GP} is the Gunn-Peterson optical depth from (3.49), W_α is the normalized asymptotic spatial distribution of Lyα photons and L_ν is the luminosity of the source expressed as number of photons per unit frequency per unit time.

S_α is a back-reaction correction from [7]. P_α affects the population in the hyperfine states, such that we can not necessarily assume that 3/4 of all neutral hydrogen is in the triplet state. If the population deviates significantly due to the spin exchange, this affects the average energy required for lyα, which again affects the colour temperature T_c and lyα scattering rate (affecting thxe coupling constant, and thus again the spin temperature). In that sense (3.48) must be seen as an implicit equation which should be solved using an iterative algorithm. In [17] the back-reaction correction due to spin exchange is studied extensively, but they find that the effects are small for $T_S, T_k \gg 1$.

Equivalent to (4.1), we may write the scattering rate for the "injected" photons as:

$$P_\alpha^{\text{inj},n}(r) = \tau_{\text{GP}} S_\alpha \int_{\nu_n}^{\nu_{n+1}} \left[\sum_{k=1}^{\infty} f_{\text{recycle},n} (1 - P_{nP\to 1S}) (P_{nP\to 1S})^{k-1} W_{k,n} \right] L_\nu \, d\nu \tag{4.2}$$

Where $f_{\text{recycle},n}$ is the recycle fraction of Lyn from (3.51) and $P_{nP\to 1S}$ is the decay probability to the ground state from (3.50), which is equivalent to the probability that a photon does not get photoconverted during a scattering. $W_{k,n}$ is the normalized spatial distribution of Lyn photons after k scatterings. As $(1 - P_{nP\to 1S})(P_{nP\to 1S})^{k-1} \to 0$ when $k \to \infty$, we only need to look at the first scatterings, consistent with our expectations from section 3.2.4.

The total scattering rate may then be written by combining the contributions from (4.1) and (4.2):

$$P_\alpha(r) = P_\alpha^{\text{cont}}(r) + \sum_{n=3}^{\infty} P_\alpha^{\text{inj},n}(r)$$

In [8] they used a simple power law for the luminosity from the source $L_\nu \propto \nu^{\alpha_s - 1}$, such that $\alpha_s = 1$ corresponds to a flat spectrum, and the values $\alpha_s = 1.29, 0.14$ have been used as approximations to PopIII and PopII star populations spectra respectively.

One of their main discoveries was that the scattering rate $P_\alpha(r)$ is much steeper than would initially expect by assuming the line profile is a dirac δ function and no photons scatter before reaching the thermal core. In fact, as the optical depth is so high they may scatter far out in the Lorentz wing of the line, and may therefore change direction while redshifting, effectively travelling a smaller distance away from the central source before resonance. This means that the naively assumed scattering rate following a $1/r^2$ drop is incorrect, and the deviation from the slope assuming a δ function increases with decreasing initial frequency offset $\Delta\nu$ such that we see a very significant deviation in the lower Lyman transitions.

We therefore rather expect the continuum photons to follow $P_\alpha^{\text{cont}} \propto r^{-7/3}$, and each individual injected photon follows $P_\alpha^{\text{inj},n} \propto r^{-5/2}$ very close to the central source and further

away as a truncated r^{-2} slope, assuming the model from [8]. The sum of all injected photons can therefore be shown to also follow $P_\alpha^{\text{inj}} \propto r^{-7/3}$, implying that the total scattering rate in the early universe might be up to 50% higher than expected.

4.2 Improving the model

In this thesis we want to explore what effects improving this model has on the scattering rate. We make the spectrum from the source more realistic than a simple power-law by adding a variable column density of atomic hydrogen, such that we get Lyman absorption lines in the spectrum. In addition we assume a fraction of H_2 in the central source with absorption lines. In subsection 4.3.3 we will go further into how this is done in the code.

We also add the possibility of having a spherically symmetric ionisation bubble around the central source where no photons may scatter. The implementation of an ionised bubble is discussed in subsection 4.3.4.

We expect UV-emitting galaxies to be surrounded by regions with high densities of hydrogen as galaxy formation is biased towards overdensities. If there is a sufficient amount of neutral hydrogen surrounding the UV-emitting source, a large fraction of the photons with frequencies near a Lyman transition will be absorbed by this medium. This means that the surrounding IGM will not recieve these photons. By assuming the spectrum is simply a powerlaw, we may neglect the frequency bias that this entails. The spectrum for atomic hydrogen that is inserted into the Monte Carlo radiative transfer code will therefore follow a shape similar to the one shown in Figure 4.1.

Even as the UV source ionises its surroundings we expect some HI causing absorption lines. Due to the high densities we should expect a high recombination rate and also possibly DLA-like structures, causing some photons to be absorbed. We assume absorbtion by atomic hydrogen should affect frequencies near ν_α and the Lyman limit ν_∞ the most. For the most massive sources, with high column densities of neutral hydrogen N_{HI} this becomes increasingly important. We can see this by comparing the input spectra from $N_{\text{HI}} = 10^{25}$ m^{-2} and $N_{\text{HI}} = 27$ m^{-2}; line saturation and very large wings makes the spectrum a poor fit to any powerlaw.

Adding the molecular lines follows the same line of motivation as atomic lines. We expect the effect to be smaller than that of atomic hydrogen, yet still important for massive sources and sources with a high fraction of H_2. Due to self-shielding and high concentration of free electrons, we expect some halos may contain a non-negligible fraction of molecular hydrogen x_{H_2} even after modest Lyman-Werner photon and ionising photon emission, even more so if we expect significant X-ray ionisation. Spectra including H_2 lines are shown in Figure 4.1.

During reionisation, ionisation bubbles start forming. As HII has no bound electrons, no Lyman transitions may occur, and therefore we expect that the scattering rate may significantly depend on the size of the ionised region. In particular this means that we will expect no injection of Lyα in this region, as the photons will be unable to get absorbed.

This also means that a photon may, given that the bubble is sufficiently large, redshift

past the thermal core of Lyn, such that it will redshift further until it scatters as a Ly$(n-1)$ photon. In addition, due to no scattering on the smallest scales, we expect photons to travel further before it's first scattering in the IGM. Both of these factors may contribute to shifting the resonant scattering out to larger scales.

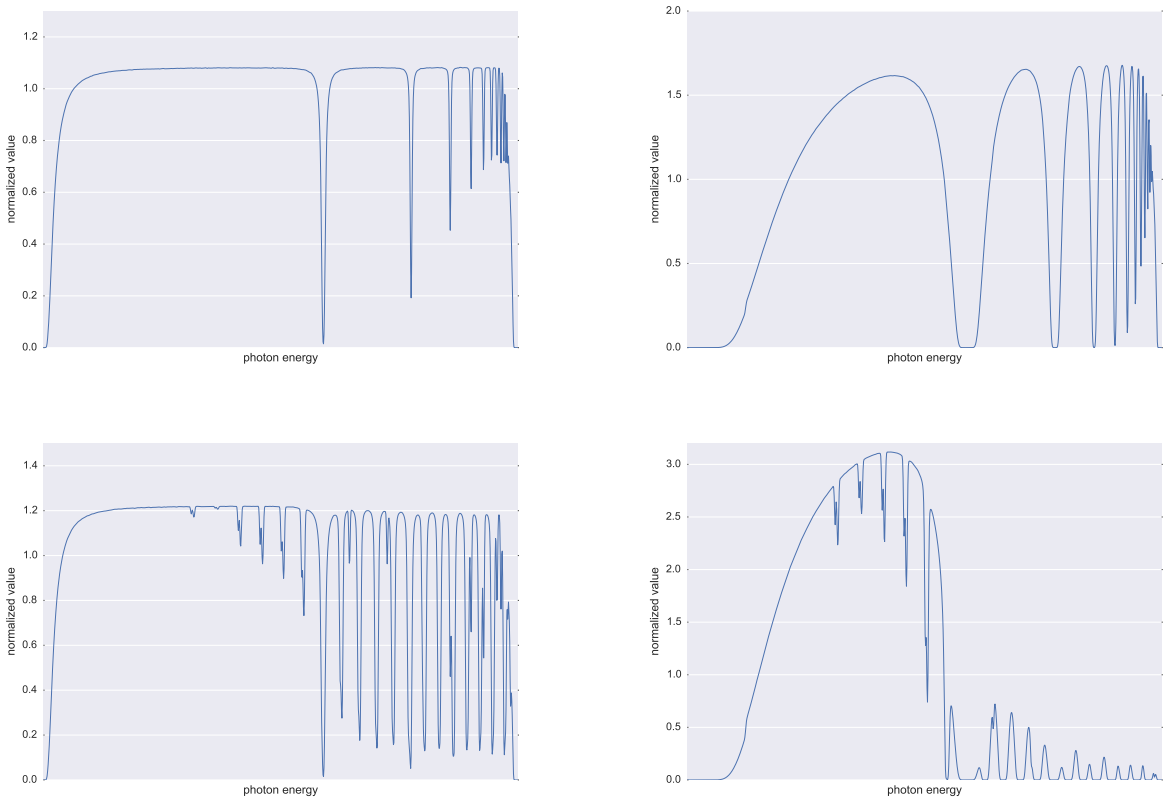

Figure 4.1: Input spectrum assuming spectral slope ν^{α_s} where $\alpha_s = 1$ between $h\nu_\alpha$ ~10.2 eV on the left and $h\nu_\infty$ ~13.6 eV on the right. The top left (top right) figure has atomic hydrogen absorption features using a column density $N_{\rm HI} = 10^{25}\,(10^{27})$ m^{-2}. The bottom left (bottom right) figure has both atomic and molecular absorption features, column density $N_{\rm HI} = 10^{25}\,(10^{27})$ m^{-2} and a molecular fraction $x_{H_2} = 10^{-2}$, which is quite high. Figures shown are normalized as $\int_{\nu_\alpha}^{\nu_\infty} L_\nu = \nu_\infty - \nu_\alpha$, such that deviations of L_ν from 1 correspond to the bias of including absorption features, compared to simply using a power-law.

4.3 Code

4.3.1 Monte Carlo radiative transfer

To model this system we in principle need to solve the transport equation (3.19), and thereby compute the intensity $I_\nu(\mathbf{r})$ throughout the IGM. Analytic solutions to the transport equation only exist for a few very simple cases. Instead, many radiative transfer

problems are handled using the Monte Carlo method, where we follow individual photons as they scatter until they decay or reach resonance. This method has the major advantage that it is very easy to modify, and is in its nature similar to conducting an experiment, making it very intuitive. It is also easy to parallelize the Monte Carlo method, if formulated such that each photon is independent of other photons. In fact, for this thesis, the code implemented uses a hybrid parallel programming model combining MPI with OpenMP. The downside of using Monte Carlo is that it converges slowly towards the true solution, and is therefore typically computationally expensive. The basic algorithm for our model is the following:

1. Start with a photon at the central source. Adopting the distance parameter $\beta = Hr/c$ from [8] we may write the initial position as $\beta_i = 0$. We also give the photon a randomly chosen frequency drawn based on the input spectrum[1] and defining a frequency offset with respect to the line centre ν_0 as $\frac{\Delta\nu}{\nu_0}$.

2. If $\nu > \nu_\beta$ (ie. not continuum photon), use rejection sampling with the recycle fraction f_{recycle} from (3.51) to determine whether it will become an injected Lyα photon. If not, go back to step 1.

3. Draw an optical depth to the next scattering according to the probability distribution $P(\tau) = e^{-\tau}$. This may be done by drawing a random uniformly distributed number p between 0 and 1, such that the randomly drawn optical depth is $\tau_p = -\ln(p)$

4. Find the distance $\Delta\beta$ travelled corresponding to τ. The scattering optical depth of the photon after having travelled a distance Δr may be written as:

$$\tau = \int n_{\text{HI}}\sigma(\nu - \nu\beta(r))\, dr = \tau_{\text{core}}\Delta\beta \left[\frac{\Delta\nu}{\nu_0}\left(\frac{\Delta\nu}{\nu_0} - \Delta\beta\right)\right]^{-1} \tag{4.3}$$

   with $\tau_{\text{core}} = n_{\text{HI}}\sigma_{\nu_0}c/H$. We can here assume that we are in the wing such that the profile is Lorenzian: $\sigma \propto \sigma_{\nu_0}\left(\frac{\Delta\nu}{\nu_0}\right)^{-2}$. Inverting (4.3) we get $\Delta\beta(\tau)$

5. Determine the new position of the photon $\boldsymbol{\beta}_f = \boldsymbol{\beta}_i + \Delta\boldsymbol{\beta}$ using a random angle θ such that $\cos\theta \in [-1, 1]$ is uniformly distributed.

6. Redshift the photon according to $\left(\frac{\Delta\nu}{\nu_0}\right)_f = \left(\frac{\Delta\nu}{\nu_0}\right)_i - \Delta\beta$ and define the initial positon for the next step: $\beta_i \leftarrow \beta_f$.

7. If the photon is not a continuum photon, check if it decays using rejection sampling with P_{if} from (3.50). If yes, skip step 8.

[1] For a flat spectrum with no absorption features this simply corresponds to drawing from a uniform distribution with freqency in $10.2\ \text{eV}/h < \nu_i < 13.6\ \text{eV}/h$, while for a power law spectrum with absorption features we use Monte Carlo rejection sampling on the distribution given by the power law.

8. Repeat 3–7 until photon decays or $\Delta\nu/\nu_0 < \Delta_{\text{core}}$, where Δ_{core} is on the order of the width of the thermal core.

This process is the repeated until the set number of photons have been drawn. Note that the angle here is uniformly chosen, as opposed to a dipole distribution, which we expect from photons. [8] tested this, but determined that its effect was small.

Rejection sampling is the simple process of comparing a value with a random number chosen by the distribution we expect the value to follow in the given circumstances. For an event $X(x', y')$ with x following a probability distribution $y(x)$ and $y' = y(x')$ we may simply compare a random number $y_R \in [y_{\min}(x'), y_{\max}(x')]$ with $y(x')$. If $y' \geq y_R$ we accept X, otherwise the event X does not occur. In the next section we will address how we convert a uniform distribution from 0 to 1 and turn this into our desired probability distribution.

4.3.2 Sampling the power law distribution

How do we sample a quantity properly given that we know the distribution? The frequencies from [8] follow a power law distribution given by $f(\nu) = \nu^{\alpha_s - 1}$. In general, we may say that the relative weighting of each frequency is given by $w_\nu = f(\nu)$. When $\alpha_s = 1$ we get $w_\nu = 1$, corresponding to frequency-independent (equal) weighting as we would expect from a uniform distribution.

We require that the corresponding probability distribution should be equal 1 when integrated, $\int P_\nu \, d\nu \equiv 1$. The probability therefore proportional to the weighting with some normalization parameter N. We then get (for $\alpha_s = 1$):

$$\int P_\nu \, d\nu = \int_{\nu_{\min}}^{\nu_{\max}} N w_\nu \, d\nu = N(\nu_{\max} - \nu_{\min})$$

Using this combined with the normalization requirement we get

$$N = \frac{1}{\nu_{\max} - \nu_{\min}} \Rightarrow P = \frac{1}{\nu_{\max} - \nu_{\min}}$$

We wish to express frequencies given by a random number p between 0 and 1 which follows this distribution. We can then express our sampled frequency ν_p as a function of the random number:

$$p = \int_{\nu_{\min}}^{\nu_p} P_\nu \, d\nu$$
$$\nu_p = \nu_{\min} + p(\nu_{\max} - \nu_{\min})$$

We can apply the same procedure to the less trivial example of $\alpha_s \neq 0$ to get:

$$\nu_p = [\nu_{\min}^{\alpha_s} + p(\nu_{\max}^{\alpha_s} - \nu_{\min}^{\alpha_s})]^{1/\alpha_s}, \tag{4.4}$$

which in the limit $\alpha_s = 1$ reduces to the uniform distribution in (4.4). For $\alpha_s = 0$ we get the following distribution:

$$\nu_p^{\alpha_s} = \nu_{\min} \left(\frac{\nu_{\max}}{\nu_{\min}} \right)^p. \tag{4.5}$$

This may then be used to compute (4.1) and (4.1) using the appropriate values for $\nu_{\min}$ and $\nu_{\max}$ given by the integration limits, once $W_{n,k}$ has been determined.

4.3.3 Absorption

While varying the powerlaw distribution effectively tilts the spectrum, absorption adds more frequency-sensitive structure. To model the absorption features we first sample the spectrum according to its powerlaw distribution, then check accept or reject the value based on evaluating the optical depth. Modelling atomic hydrogen lines and molecular hydrogen lines is in principle the same once you have obtained the relevant Einstein coefficients and oscillator strengths.

We assume the line has a Voigt profile from (3.30), which we for efficiency simplify in the following way:

$$\phi(\nu) \approx \begin{cases} e^{-x^2} & |x| < x_{\text{crit}} \\ \frac{a}{\sqrt{\pi}} \frac{1}{x^2} & |x| > x_{\text{crit}} \end{cases} \tag{4.6}$$

In the line core the line is approximated as a Gaussian while in the wing where the exponential becomes subdominant, it is approximated as $\propto x^{-2}$. The transition between these two regions is around $x_{\text{crit}} = 3.2$ as done in [10]. The Voigt profile approximation from [27] is also implemented. This is more accurate, in particular around the transition between the core and the wing, but as the simple approximation using (4.6) is sufficient for our model, it is preferred due to efficiency.

The line profiles for each line is then evaluated. Using this we approximate the optical depth as:

$$\tau_\nu = \sigma_{\text{core}} N_{\text{HI}} \phi(\nu)$$

where N_{HI} is the column density of neutral hydrogen, and σ_{core} is the cross section at line centre computed in (B). If the line is a molecular hydrogen line, we multiply this expression by x_{H_2}. From this we may evaluate the expression $p < e^{-\tau}$. If this holds true for all lines, the photon is accepted, otherwise it is rejected and we resample the spectrum.

The energy levels, Einstein coefficients and oscillator strengths of the atomic hydrogen are all computed from scratch in my code, while for the molecular hydrogen tabulated values are used.

4.3.4 Ionised bubble

In implementing an ionised bubble, all one needs to do is prevent the photons from scattering while inside this region. While this is conceptually and computationally simple, there are some techicalities involved.

Computing the optical depth (ie. travel distance) is based on the IGM having a uniform density given by the global average, (3.24), with $x_{HI} = 1$. In the ionised bubble this final assumption is obviously wrong. To solve this without having to compute the ionised fraction along the optical path, the first scatter will always occur after an initial purely radial displacement which we may implement by changing the initial position to the distance corresponding to the Strömgren radius: $\beta_i = \beta_{R_s}$, and redshifting it correspondingly.

If the photon later scatters back into the bubble, however, it is less simple. The photon may, depending upon the direction it travels, travel straight across the bubble, $|\Delta\boldsymbol{\beta}| \sim 2\beta_{R_s}$, or travel along the rim of the bubble, $|\Delta\boldsymbol{\beta}| < \beta_{R_s}$. Either way we still use a optical depth which assumes uniform number density of neutral hydrogen. This is acceptable as this is chosen randomly, as long as we do not treat scatterings within the ionised bubble as actual scatterings. This means that we need a way to keep propagating the photon in the same direction if such a "pseudo-scattering" occurs.

Assuming after a number of scatterings a photon scatters from a position β_1 with the angle θ, where $\theta = 0$ corresponds to travelling radially away from the central source as shown in Figure 4.2. If it after travelling a distance $\Delta\beta$ to a position β_2 withing the bubble, it should (pseudo)scatter such that it continues in the same direction, defined as θ' with respect to the central source. If we define the difference between these two angles as seen radially, as $\alpha \equiv \theta - \theta'$ it should propagate with the angle:

$$\mu_{\theta'} = \mu_\theta \mu_\alpha + \sqrt{(1-\mu_\theta^2)(1-\mu_\alpha^2)}, \quad \mu_\alpha = \frac{\beta_1^2 + \beta_2^2 - \Delta\beta^2}{2\beta_1\beta_2},$$

where $\mu_i \equiv \cos i$. We add a travel distance $\Delta\beta'$ corresponding to the remaining path within the bubble in addition to the normal distance determined by the optical depth:

$$\Delta\beta' = \beta_2\mu_{\theta'}\left(1 \pm \sqrt{1 + \frac{\beta_{R_s}^2 - \beta_2^2}{\beta_2^2\mu_{\theta'}^2}}\right) + \frac{\Delta\nu}{\nu}\left(1 - \frac{1}{1 + \frac{\Delta\nu}{\nu}\frac{\tau}{\tau_{\text{core}}}}\right)$$

where the sign is determined by the sign of $-\mu_{\theta'}$. This physically corresponds to whether the photon is heading towards the central source or not at the point of the pseudo-scatter.

Finally, we need to consider the case where the photon redshifts past the thermal core of the Lyman transition. What we naively would assume would become a Lyn photon will then typically first be injected into Lyman resonance when it approaches the thermal core of Ly$(n-1)$. For $n > 1$ we then simply redefine the offset with respect to the next transition, while for Lyα it will no longer have sufficent energy to excite hydrogen to the $2P$ state.

Although this is physically correct this will bias our model negatively so we will get no such skipping of the line core of Lyβ to Lyα, as we do not model these due to them not contributing to P_α. The remedy we have chosen to implement is to assume the loss/-gain of photons in one injected photon population due to this effect is similar to that of its neighbours, $\delta\text{Ly}n \sim \delta\text{Ly}(n-1)$, which should hold when assuming small ionisation bubbles, similar transition profiles, and for similar production rates (i.e., if the source has

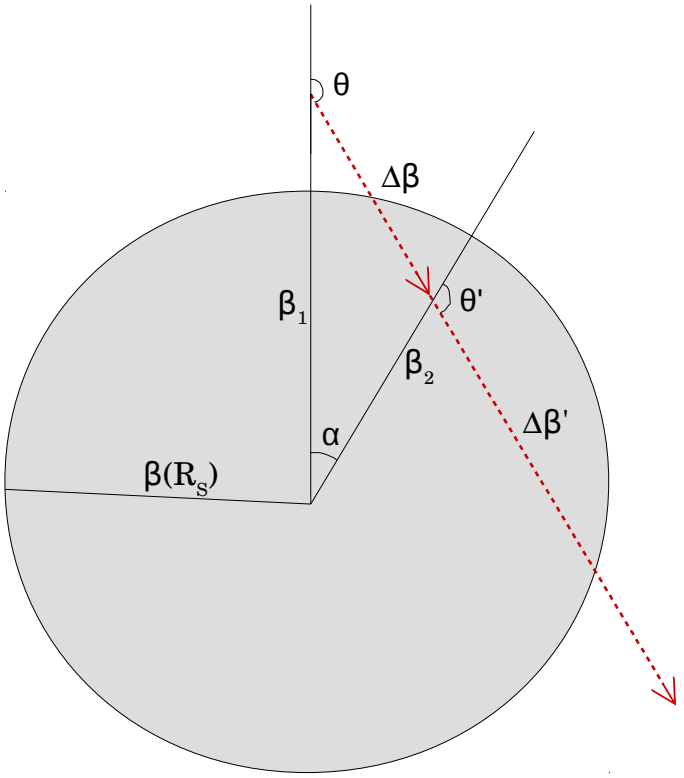

Figure 4.2: Illustration of a "pseudo-scattering" scenario. Photon starts outside the ionised bubble which has the size β_{R_s} at a distance β_1 from the central source, where it scatters and radiates into the ionised bubble to β_2, traversing $\Delta\beta$. The scattering angle θ needs to be transformed into θ' to preserve the direction of the photon while it is travelling through the ionised bubble. The photon then travels a distance $\Delta\beta'$ before scattering outside the bubble.

a reasonably flat spectrum). As we do not model Lyβ, we may assume the loss of Lyγ photons due to this effect is equal to the gain of Lyα. This should be a marginal effect however, and is only added for completeness.

Chapter 5

Results

5.1 Reproducing previous results

In [8] they show the effect of modelling the wing scattering of the Lyman transitions. In Figure 5.1 we see what effect this has on Lyα at a given frequency offset $\Delta\nu/\nu$. Previously one assumed no wing scattering, such that the photons simply redshift radially out to a distance corresponding to the offset. This is shown to be a poor assumption, in particular for smaller offsets of Lyα. They show that for $\Delta\nu/\nu = 10^{-2}$ an average photon travels $\sim 45\%$ less than previously assumed, and for $\Delta\nu/\nu = 10^{-3}$ the distribution peaks at $\sim 15\%$ of what one would expect by assuming only core scattering, and thus a dirac profile δ_D distribution.From Figure 5.1 we also see how the distribution approaches its asymptotic distribution after $N_{\text{scatter,wing}} \sim 10^2$. Only using wing scattering to determine the asymptotic distribution; and then only using core scattering $N_{\text{scatter,core}} \sim 10^6 \gg N_{\text{scatter,wing}}$ to determine the scattering rate, and thus the spin temperature, is therefore a reasonable approximation.

Looking at the radial probabilty distribution for injected photons [8] find that wing scattering is less important for the higher transitions, primarily due to a smaller cross-section. Thus the δ_D function is a more reasonable approximation for these photons, as we can see from Figure 5.2. This, combined with the possibility of decaying to a injected Lyα at every scattering contribute to the distribution reaching its asymptotic form after $N_{\text{scatter,wing}} \sim 10$, less scatterings than the "continuum" photons.

For both Figure 5.1 and Figure 5.2 any deviations between the results reproduced here and the ones in [8] are negligible, and primarily caused by using a finite amount of photons, and the models using different cosmological parameters.

In Figure 5.3 we see slight deviations on small scales, which can be removed by using more lyman transitions (For this model 50 Lyman transitions are used, while in [8] higher Lyman transitions are approximated by extrapolation). Otherwise we find good agreement with previous results. Figure 5.3 show the radial scattering rate profile from a central source, assuming uniform frequency distribution. We see that "injected" photons play an important role on small scales, but on scales larger than $\sim 5\,\text{h}^{-1}$ Mpc comoving, the main

contribution is from the "continuum" photons. Each of the "injected" Lyman follow an approximate $1/r^2$ slope truncated at $Hr/c = (\nu_{n+1} - \nu_n)/\nu_n$, corresponding to the larges possible initial frequency offset. This slope reflects what was shown in Figure 5.2; the δ_D function is reasonable. The combined contribution of all the "injected" photons add upp to a slope $\sum_n P_\alpha^{\mathrm{inj},n} \propto r^{-7/3}$.

We get the same slope of the scattering rate for the "continuum" photons, but for different reasons. Assuming wing scattering the mean free path goes as $l \propto (\Delta\nu)^2$, while the number of scatterings goes as $N_{\mathrm{scatter}} \propto \Delta\nu/l$, and if it scatters randomly it should evolve similar to a random walk with step size l per scattering, $r \propto N_{\mathrm{scatter}}^{1/2} l$. Combining this, within a volume $V \sim 1/r^3$ we therefore expect $P_\alpha \propto \Delta\nu/r^3 \propto r^{-7/3}$ as mentioned in [8]. From the bottom panels of Figure 5.3 we see that for all scales but the very largest, this corresponds to an increase in the overall scattering rate by $\sim 40\%$ compared to earlier calculations, which may change by 5% when looking at reasonable power-law spectrum slopes.

In Figure 5.4 the we look at the effects of variation of the spectrum slope. By having a harder slope we, perhaps not surprisingly, see that contributions from the "injected" photons increase, while the opposite is true for the "continuum" photons, except for on the largest scales, as these correspond to the most energetic "continuum" photons (ie. with frequency offset $\lesssim (\nu_\beta - \nu_\alpha)/\nu_\alpha$). We also see that the "no wing scattering" assumption will tend to make the spectrum more sensitive to the slope.

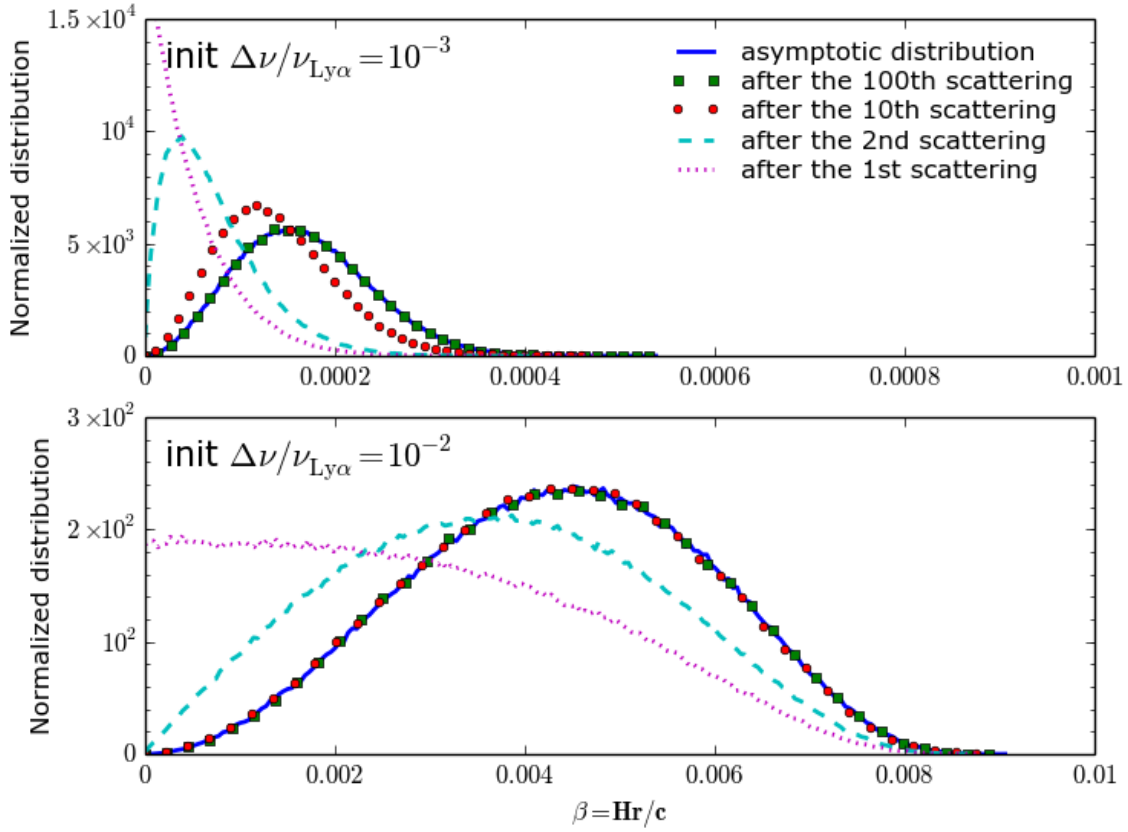

Figure 5.1: Radial probability distribution of "continuum" Lyα photons, $4\pi\beta^2 W_{k,\alpha}$, around central source at $z = 25$. In the plots the photons start with a frequency offset of $\Delta\nu/\nu_\alpha$. Previous calculations assume a δ_D distribution at $Hr/c = \Delta\nu/\nu$, corresponding to the value on the right of the plots, implying a 1:1 correspondence between offset and resonance position, which here becomes a distribution. In the top (bottom) plot the initial frequency offset of the photons are set to $\Delta\nu/\nu_\alpha = 10^{-3}\,(10^{-2})$. The different curves show the distribution after a given amount of scatterings, with the solid blue curve corresponding to the asymptotic distribution these photons would have while in line resonance.

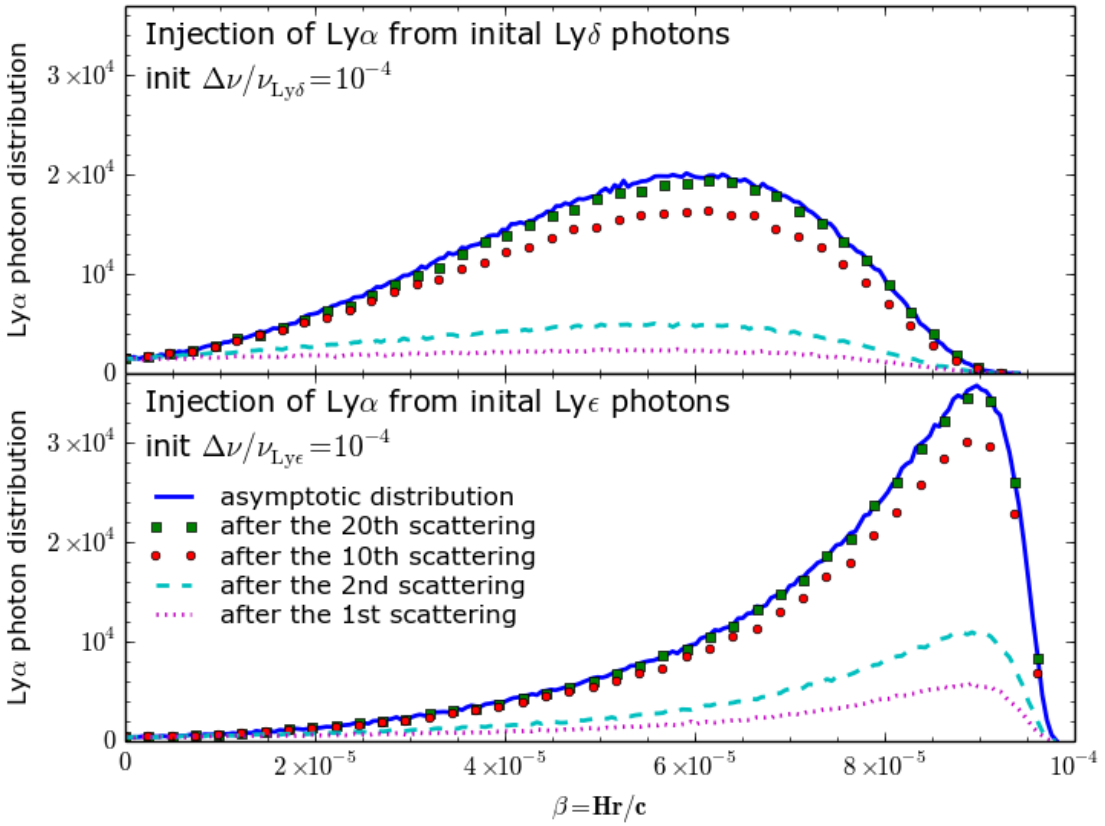

Figure 5.2: Radial probability distribution of "injected" Lyα photons, $4\pi\beta^2 W_{k,n}$, around central source at $z = 25$. In the plots the photons start with a frequency offset of $\Delta\nu/\nu_n$. In the top (bottom) plot Ly$\delta(\epsilon)$ photon distribution given an initial frequency offset of the photons at $\Delta\nu/\nu_\alpha = 10^{-4}$. The different curves show the distribution after a given amount of scatterings, with the solid blue curve corresponding to the asymptotic distribution these photons would have while in line resonance.

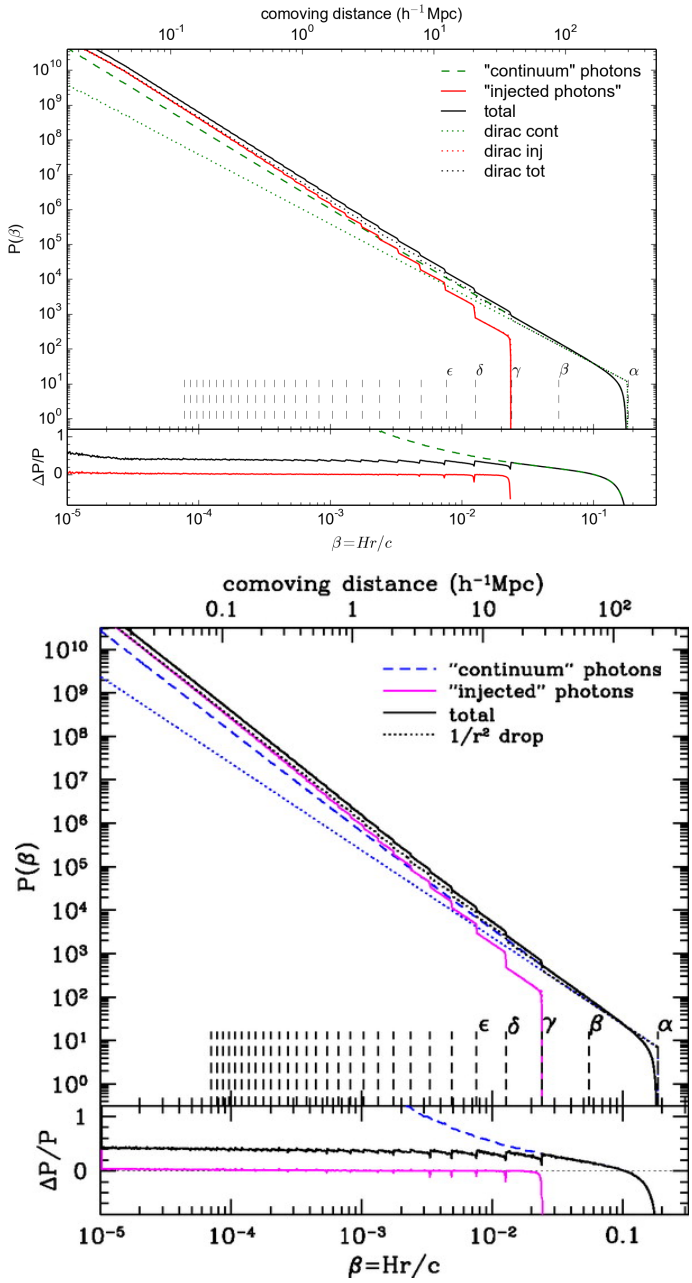

Figure 5.3: Radial Lyα scattering rate profile around central source at $z = 25$. The profile assumes a flat UV spectrum taking into account wing scattering (solid/striped lines) is compared to the results we get by assuming no wing scattering (dotted lines), eg. a δ_D distribution for a any inital frequency offset. The black curves correspond to the total scattering rate, green/blue curves only include contributions from the "continuum" photons while red/magenta curves correspond to the "injected photons". The vertical striped lines show the maximum distance a photon of a given Lyman transition may reach. In the bottom panel the fractional difference with respect to its δ_D counterparts are shown. The top plot is made using the model used for this thesis in an attempt to reproduce the results from the bottom plot, which is from [8].

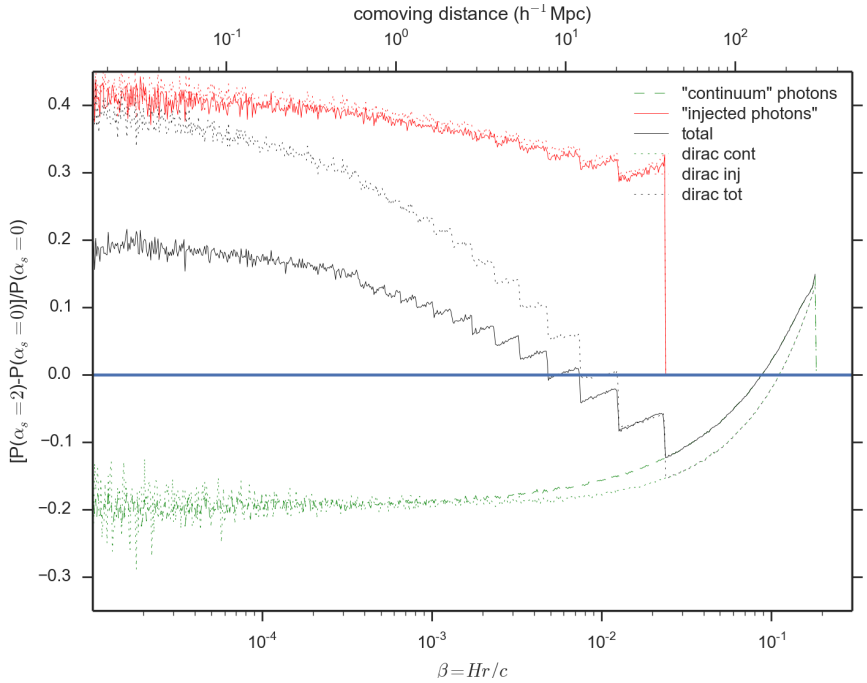

Figure 5.4: Fractional difference in scattering rate profile for different slopes at $z = 25$. The input spectrum from the central source is normalized so that an equal amount of photons are emitted from the spectra. The profiles follow a power law $\propto \nu_s^{\alpha}$, with $\alpha_s = 2$ and $\alpha_s = 0$. This plot is particularly susceptible to uncertainties in the Monte Carlo method as it compares the differences beween two very similar curves.

5.2 Model extension results

When adding atomic absorption lines to the input spectrum, this primarily affects the population Lyα photons with very small initial frequency offsets, and the Lyman transitions near the Lyman limit, as we can see from the top two plots in Figure 4.1. There is a sharp drop near $2{\cdot}10^{-1}\,h^{-1}$ Mpc, where the absorption lines near the Lyman limit ν_∞ in the input spectrum are so close that they overlap, and there is almost no emission. The Lyα line is saturated at the relevant input spectrum column densities, causing a drop in luminosity on the largest scales.

The luminosity at the rest of the frequencies is thus effectively boosted relative to the average. For larger column densities when Lyβ and Lyγ and Lyδ saturate, this decreases the amount of photons from this frequency range, decreasing the scattering rate in the range 10–$10^2\,h^{-1}$ Mpc up to a factor of ~ 5 between $N_{\text{HI}} = 10^{25}$ m^{-2} and $N_{\text{HI}} = 10^{25}$ m^{-2}. The atomic absorption features therefore highlight the contribution from "continuum" photons with high initial frequency offset.

Adding the molecular lines to the input spectrum adds an additional suppression of Lyn photons for $n > 2$. Although not as significant as changing the column density, this also contributes to the relative boost of the "continuum" photons, as can be seen from Figure 5.6.

Adding an ionised bubble around the central source prevent any photons from scattering there. This will cause the highest Lyman transitions to be skipped, as they all have very small initial frequency offsets. We therefore see a peak in the scattering rate profile right outside the bubble as the photons which already are in their thermal core may scatter and decay. To a lesser extent this will also happen to less energetic Lyman transitions which may scatter repeatedly across the bubble, redshifting into resonance after fewer wing scatterings. Interestingly this also causes peaks near each of the profile cores for the same reason. These peaks are pushed towards lower transitions for larger bubble sizes.

Combining these effects we get Figure 5.8. We see that "continuum" Lyα photons are strongly suppressed on small scales due to the Lyα absorption line. The highest Lyman transitions are also suppressed on scales $\beta \sim \beta_{R_s}$, both due to the ionisation bubble, but also due to the input spectrum from the source has absorbed most of the Lyman limit photons. The effect is that even though the "continuum" photons are subdominant compared to "injected photons" for a significantly larger frequency range than in Figure 5.3, by up to several orders of magnitude at smaller scales. The net effect on "continuum" photons on larger scales is positive overall when adding atomic absorption and molecular absorption or ionised bubbles separately, and so the combined effect is non-negligible. These effects push the scatting to larger scales, which ultimately implies that the coupling of the spin temperature to the Lyα colour temperature rate should occur later than previously assumed on large scales.

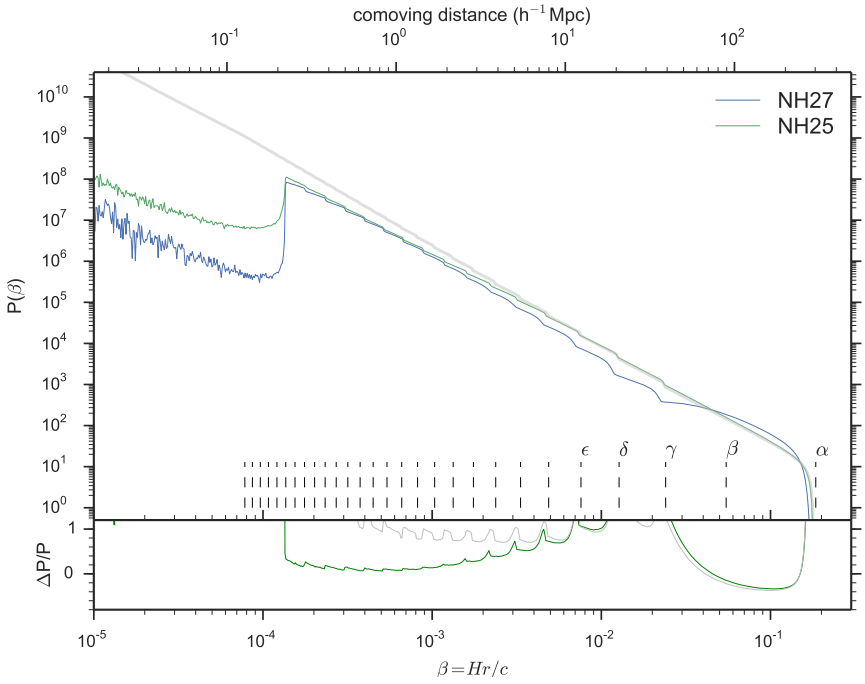

Figure 5.5: Radial Lyα scattering rate profile around central source at $z = 25$. The profile assumes a UV spectrum expected by PopII stars, although the effect of changing this is not to significant. The blue(green) curve corresponds to the total scattering rate given a column density $N_{\rm HI} = 10^{27}(10^{25})$ m^{-2}, while the gray corresponds to the total scattering rate assuming the flat UV spectrum with wing scattering(ie. black curve from Figure 5.3). The vertical striped lines show the maximum distance a photon of a given Lyman transition may reach. In the bottom panel the fractional difference between $N_{\rm HI} = 10^{27}$ m^{-2} and the other curves are shown in green($N_{\rm HI} = 10^{25}$ Mpc) and gray(flat UV spectrum).

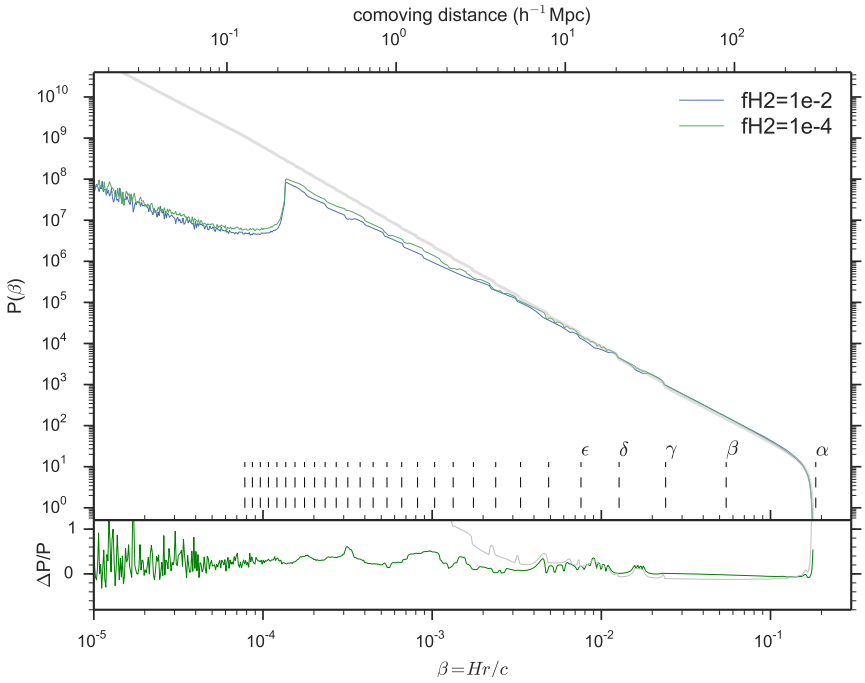

Figure 5.6: Radial Lyα scattering rate profile around central source at $z = 25$. The profile assumes a UV spectrum expected by PopII stars, although the effect of changing this is not to significant. The blue(green) curve corresponds to the total scattering rate given a column density $N_{HI} = 10^{25}$ m^{-2} and molecular hydrogen fractions $x_{H_2} = 10^{-2}(10^{-4})$(denoted fH2), while the gray corresponds to the total scattering rate assuming the flat UV spectrum with wing scattering(ie. black curve from Figure 5.3). The vertical striped lines show the maximum distance a photon of a given Lyman transition may reach. In the bottom panel the fractional difference between $x_{H_2} = 10^{-2}$ and the other curves are shown in green($x_{H_2} = 10^{-4}$ Mpc) and gray(flat UV spectrum).

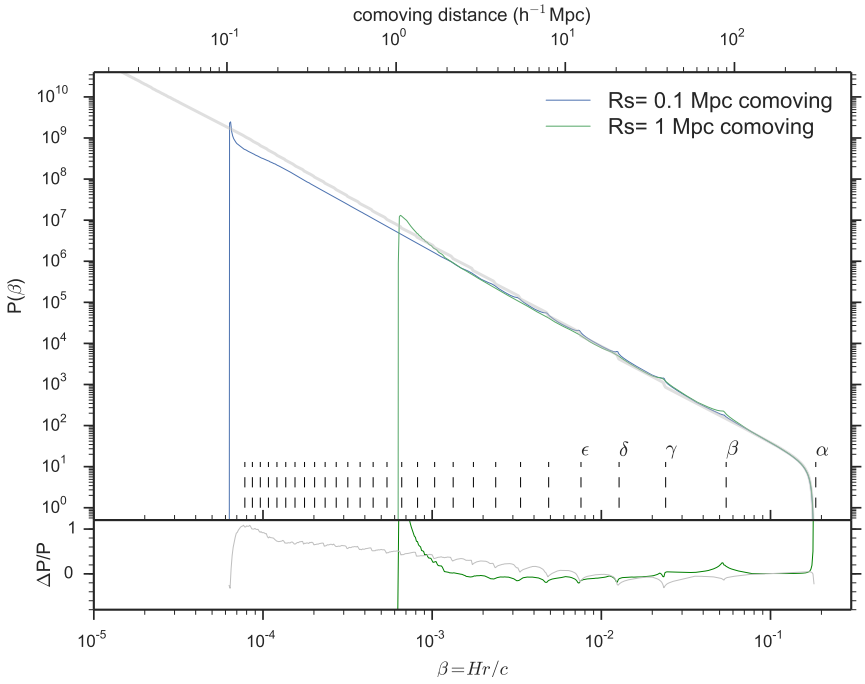

Figure 5.7: Radial Lyα scattering rate profile around central source at $z = 25$. The profile assumes a UV spectrum expected by PopII stars, although the effect of changing this is not to significant. The blue(green) curve corresponds to the total scattering rate given an ionised bubble $\beta_{R_s} = 0.1(1.0)\,h^{-1}$ Mpc, while the gray corresponds to the total scattering rate assuming the flat UV spectrum with wing scattering(ie. black curve from Figure 5.3). The vertical striped lines show the maximum distance a photon of a given Lyman transition may reach. In the bottom panel the fractional difference between $\beta(R_s) = 0.1\,h^{-1}$ Mpc and the curves are shown in green($\beta_{R_s} = 0.1\,h^{-1}$ Mpc) and gray(flat UV spectrum).

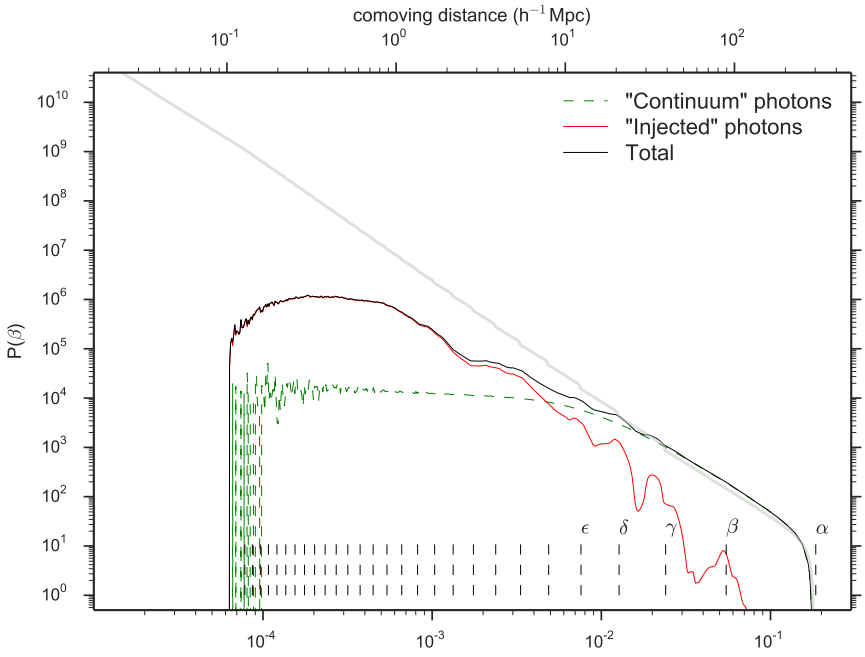

Figure 5.8: Radial Lyα scattering rate profile around central source at $z = 25$. The profile assumes a UV spectrum expected by PopII stars, although the effect of changing this is not to significant. The blue(green) curve corresponds to the total scattering rate given an ionised bubble $\beta_{R_s} = 0.1(1.0)\,h^{-1}$ Mpc, while the gray corresponds to the total scattering rate assuming the flat UV spectrum with wing scattering(ie. black curve from Figure 5.3). The vertical striped lines show the maximum distance a photon of a given Lyman transition may reach. We see that overall the photons are pushed from smaller scales to larger, compared to the results from [8]

Chapter 6

Conclusions & Outlook

In this thesis we have implemented a basic radiative transfer code to model the radial Lyα scattering rate around the epoch when the first stars and galaxies start forming, and ionizing their surroundings. The model, based on a model by [8], additionally includes absorption lines from atomic and molecular hydrogen, and ionizing bubbles. We have shown that these will suppress scattering on small scales, and thus enhance the relative scattering rate on scales of 10 – $10^2\,h^{-1}$ Mpc comoving.

In particular, the effects of saturated higher order Lyman transitions in the source will push the scattering rate to larger scales. "Continuum" photons with frequency offsets very near $(\nu_\beta - \nu_\alpha)/\nu_\alpha$, corresponding to the photons with the largest photon offsets, will be suppressed due to the very large equivalent width of Lyβ absorption by interstellar HI in the central source. This is a marginal effect, dependent on large column widths. "Injected" photons will dominate on smaller scales, where the large equivalent width of Lyβ suppresses emission of "continuum" photons. On the scales of $\lesssim 1\,h^{-1}$ Mpc contributions from "continuum" photons are several orders of magnitude smaller than the total, and these may therefore safely be neglected on this range if the column density is $N_{\rm HI} > 10^{25}$ m^{-2} $= 10^{21}$ cm^{-2}. "Injected" photons near the Lyman limit are suppressed on scales $\lesssim 0.2\,h^{-1}$ Mpc.

Molecular absorption, while a smaller effect than atomic absorption reduces the amount of "injected" photons. This may be important for sources with a large molecular hydrogen fraction x_{H_2}.

Ionised bubbles, as suspected, will push the photons out to larger scales. This will bring the spatial distribution of Lyα closer to the naively expected δ_D distribution for large ionised bubbles. This will also cause photons to skip lyman transition cores, which also brings more photons onto larger scales.

When comparing the total scattering rate profile to the one of [8], the net effect is that the scattering rate on smaller scales($< 10\,h^{-1}$ Mpc) is many orders of magnitude less, while there is an increase of photons by $\sim 20\%$ on larger scales($\sim 10^2\,h^{-1}$ Mpc) for the model for the model with parameters listed in Appendix A.

Understanding the scattering rate profile will be important in extracting information about the first galaxies and stars from the 21cm experiments which will have its golden

age in the next decades. How early the Wouthuysen-Field coupling starts will determine when the differential brightness temperature is observable.

We have shown that making a more realistic model of the scattering rate of Lyα not simply improves the results from earlier, but drastically changes the profile by several orders of magnitude. These results indicate that around the first galaxies in the dark ages, when the surrounding gas can only see very few sources, the scattering rate is much weaker than previous models suggest on smaller scales. This could point towards a more uniform heating of the IGM, as opposed to heating only the closest regions around a source. In future work we will explore the implications this has on the evolution of the spin temperature in the dark ages.

During later stages a gas element will receive photons from sources within the past light-cone, which we would have to take into account. We expect the effect to become weaker in the later stages of the dark ages, when distant sources will contribute to the gas surrounding a given source.

Bibliography

[1] Bethe H. A. and Salpeter E. E. *Quantum Mechanics of One- and Two-Electron Atoms*. New York: Academic Press, 1957.

[2] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. 1972.

[3] R. Barkana and A. Loeb. Detecting the Earliest Galaxies through Two New Sources of 21 Centimeter Fluctuations. *Astrophysical Journal*, 626:1–11, June 2005.

[4] Rennan Barkana and Abraham Loeb. In the beginning: The first sources of light and the reionization of the universe. In *Physics reports*, volume 349. Elsevier, 2001.

[5] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, J. Caruana, B. Holwerda, R. Smit, and S. Wilkins. Cosmic Reionization after Planck: The Derived Growth of the Ionizing Background now matches the Growth of the Galaxy UV Luminosity Density. *ArXiv e-prints*, March 2015.

[6] X. Chen and J. Miralda-Escudé. The Spin-Kinetic Temperature Coupling and the Heating Rate due to Lyα Scattering before Reionization: Predictions for 21 Centimeter Emission and Absorption. *Astrophysical Journal*, 602:1–11, February 2004.

[7] L. Chuzhoy and P. R. Shapiro. Ultraviolet Pumping of Hyperfine Transitions in the Light Elements, with Application to 21 cm Hydrogen and 92 cm Deuterium Lines from the Early Universe. *Astrophysical Journal*, 651:1–7, November 2006.

[8] Leonid Chuzhoy and Zheng Zheng. Radiative transfer effect on ultraviolet pumping of the 21cm line in the high-redshift universe. *Astrophysical Journal*, pages 912–918, 2007.

[9] P. C. Clark, S. C. O. Glover, R. J. Smith, T. H. Greif, R. S. Klessen, and V. Bromm. The Formation and Fragmentation of Disks Around Primordial Protostars. *Science*, 331:1040, February 2011.

[10] M. Dijkstra. Lyα Emitting Galaxies as a Probe of Reionisation. *Publications of the Astronomical Society of Australia*, 31:40, October 2014.

[11] M. Dijkstra and A. Loeb. Requirements for cosmological 21-cm masers. *New Astronomy*, 13:395–404, August 2008.

[12] Scott Dodelson. *Modern Cosmology*. Elsevier, 2003.

[13] G. B. Field. Excitation of the Hydrogen 21-CM Line. *Proceedings of the IRE*, 46:240–250, January 1958.

[14] G. B. Field. The Spin Temperature of Intergalactic Neutral Hydrogen. *Astrophysical Journal*, 129:536, May 1959.

[15] J. Franco, G. Tenorio-Tagle, and P. Bodenheimer. On the formation and expansion of H II regions. *Astrophysical Journal*, 349:126–140, January 1990.

[16] Alexandr Friedman. *On the Curvature of Space*, 1922.

[17] S. R. Furlanetto and J. R. Pritchard. The scattering of Lyman-series photons in the intergalactic medium. *MNRAS*, 372:1093–1103, November 2006.

[18] S. R. Furlanetto, M. Zaldarriaga, and L. Hernquist. The Growth of H II Regions During Reionization. *Astrophysical Journal*, 613:1–15, September 2004.

[19] Steven R. Furlanetto, S. Peng OH, and Frank H. Briggs. Cosmology at low frequencies: The 21cm transition and the high-redshift universe. In *Physics reports*, volume 433. Elsevier, 2006.

[20] W. Gordon. Zur Berechnung der Matrizen beim Wasserstoffatom. *Annalen der Physik*, 394:1031–1056, 1929.

[21] L. C. Green, P. P. Rush, and C. D. Chandler. Oscillator Strengths and Matrix Elements for the Electric Dipole Moment for Hydrogen. *Astrophysical Journal Supplement*, 3:37, April 1957.

[22] David J. Griffiths. *Introduction to quantum mechanics*. Pearson Education, second edition, 2005.

[23] I. I. Guseinov and B. A. Mamedov. Unified treatment of electric multipole transition radial matrix elements using radial parts of STO and complete orthonormal sets of ψ^{alpha}-ETO. *Radiation Physics and Chemistry*, 81:776–779, July 2012.

[24] Z. Haiman, M. J. Rees, and A. Loeb. Destruction of Molecular Hydrogen during Cosmological Reionization. *Astrophysical Journal*, 476:458–463, February 1997.

[25] C. M. Hirata. Wouthuysen-Field coupling strength and application to high-redshift 21-cm radiation. *MNRAS*, 367:259–274, March 2006.

[26] D. Hoang-Binh. An exact calculation of hydrogenic radial integrals and oscillator strengths, for principal quantum numbers up to N equal to about 1000. *Astronomy and Astrophysics*, 238:449–451, November 1990.

[27] J. Humlícek. Optimized computation of the Voigt and complex probability functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27:437–444, April 1982.

[28] A. A. Kaurov and N. Y. Gnedin. Recombination Clumping Factor during Cosmic Reionization. *Astrophysical Journal*, 787:146, June 2014.

[29] P. Laursen. Lyman {\alpha} radiative transfer in the high-redshift, dusty Universe. *ArXiv e-prints*, December 2010.

[30] A. Loeb. Probing the Universe after Cosmological Recombination through the Effect of Neutral Lithium on the Microwave Background Anisotropies. *Astrophysical Journal Letters*, 555:L1–L5, July 2001.

[31] Abraham Loeb and Steven R. Furlanetto. *The First Galaxies in the Universe*. Princeton University Press, 2013.

[32] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal newtonian gauges. *Astrophysical Journal*, pages 7–25, 1995.

[33] P. Madau, F. Haardt, and M. J. Rees. Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources. *Astrophysical Journal*, 514:648–659, April 1999.

[34] Ian D. McGreer, Andrei Mesinger, and Valentina D'Odorico. Model-independent evidence in favour of an end of reionization by $z \approx 6$. *MNRAS*, 447:499–505, 2015.

[35] A. Meiksin. Detecting the Epoch of First Light in 21-CM Radiation. In M. P. van Haarlem, editor, *Perspectives on Radio Astronomy: Science with Large Antenna Arrays*, page 37, 2000.

[36] Jordi Miralda-Escudé. The dark age of the universe. *Science*, 300, 2003.

[37] S. Naoz, S. Noter, and R. Barkana. The first stars in the universe. *MNRAS*, 373, 2006.

[38] Aravind Natarajan and Naoki Yoshida. The dark ages of the universe and hydrogen reionization. *arXiv*, 1206:0267, 2003.

[39] S. P. Oh and Z. Haiman. Second-Generation Objects in the Universe: Radiative Cooling and Collapse of Halos with Virial Temperatures above 10^4 K. *Astrophysical Journal*, 569:558–572, April 2002.

[40] J. A. Peacock. Large-scale surveys and cosmic structure. 2003.

[41] J. R. Pritchard and S. R. Furlanetto. Descending from on high: Lyman-series cascades and spin-kinetic temperature coupling in the 21-cm line. *MNRAS*, 367:1057–1066, April 2006.

[42] J. R. Pritchard and A. Loeb. Constraining the unexplored period between the dark ages and reionization with observations of the global 21 cm signal. *Physical Review D*, 82(2):023006, July 2010.

[43] R.J. Rutten. *The Generation and Transport of Radiation.* Utrecht, fourth edition, 1995.

[44] George B. Rybicki. *Radiative processes in astrophysics.* John Wiley & Sons, Inc, 1979.

[45] I. I. Sobel'Man. *An introduction to the theory of atomic spectra.* 1973.

[46] T. P. Stecher and D. A. Williams. Photodestruction of Hydrogen Molecules in H I Regions. *Astrophysical Journal*, 149:L29, July 1967.

[47] R. M. Thomas and S. Zaroubi. On the spin-temperature evolution during the epoch of reionization. *MNRAS*, 410:1377–1390, January 2011.

[48] R. M. Thomas, S. Zaroubi, B. Ciardi, A. H. Pawlik, P. Labropoulos, V. Jelić, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, G. J. A. Harker, L. V. E. Koopmans, G. Mellema, V. N. Pandey, J. Schaye, and S. Yatawatta. Fast large-scale reionization simulations. *MNRAS*, 393:32–48, February 2009.

[49] H. Y. Trac and N. Y. Gnedin. Computer Simulations of Cosmic Reionization. *Advanced Science Letters*, 4:228–243, February 2011.

[50] O. Zahn, C. L. Reichardt, L. Shaw, A. Lidz, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, O. Doré, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, and R. Williamson. Cosmic Microwave Background Constraints on the Duration and Timing of Reionization from the South Pole Telescope. *Astrophysical Journal*, 756:65, September 2012.

[51] Saleem Zaroubi. The epoch of reionization. *arXiv*, 1206:0267, 2013.

Appendix A

Model parameters

Physical constants used in the code are defined using the values given from `http://physics.nist.gov/cuu/Constants/`.

Below are the parameters used in the final combined model yielding Figure 5.8.

Parameter	value	description	notes
z	25	redshift	
Y	0.24	Helium mass fraction	
μ	$\frac{1}{(1-Y)+\frac{Y}{4}} = 1.22$	mean molecular weight	neutral medium
Ω_m	0.26	dark matter density parameter	
Ω_b	0.044	baryon density parameter	
Ω_Λ	0.74	cosmological constant density parameter	
Ω_r	$5.04 \cdot 10^{-5}$	radiation density parameter	
h	0.74	dimensionless hubble parameter	
n_{HI}	$\Omega_b(1+z)^3 \left(\frac{1-Y}{m_p}\right) \rho_{\mathrm{cr}}\,\mathrm{m}^{-3}$	neutral hydrogen number density	
σ_{core}	$2 \cdot 10^{-6}$	Thermal core width in IGM	
N_{HI}	$10^{26}\,\mathrm{m}^{-2}$	Column density of source	
T_{source}	10^4 K	Temperature of source	Virial temperature for atomic cooling
α_s	0.14	Spectrum power law slope parameter	expected slope of PopII
R_s	0.1 Mpc	Strömgren radius	reasonable values at $z = 25$: < 1 Mpc comoving
f_{H_2}	$2 \cdot 10^{-3}$	Molecular hydrogen fraction	
N_{Lyman}	50	Lyman transitions	
$N_{H_2\ \mathrm{lines}}$	76	Molecular hydorgen transitions	
N_{photons}	10^9	Number of photons in run	Only photons which contribute to P_α included.
$N_{scatter,max}$	10^2	Scatterings before assuming asymptotic value	Number is reset if photon skips resonance. Pseudo-scatters neglected.

Appendix B

Computing the Einstein coefficients

B.1 Dipole moment

A transition is in general expressed as $M_{if} = \langle \psi_f | V | \psi_i \rangle$ where V is the operator representing the physical interaction coupling the states. As the electromagnetic field couples electrons with photons, this is what determines the operator. The dipole moment between to states from (3.5) can be expressed using the radial wave functions P_{nl} such that the dipole moment can be written on the form:

$$M^2_{n,l\to n',l'} = \int_0^\infty P_{nl} r^3 P_{n'l'}\, dr$$

Incidentally, we may compute the quadrupole and octupole by replacing r^3 with r^4 or r^5 in the integral, respectively. [20] derived an exact expression for this integral:

$$\begin{aligned} M^2_{n,l\to n',l-1} =& \frac{(-1)^{n'-l}}{4(2l-1)!}\sqrt{\frac{(n+l)!(n'+l-1)!}{(n-l-1)!(n'-l)!}}\frac{(4nn')^{l+1}(n-n')^{n+n'-2l-2}}{(n+n')^{n+n'}} \\ &\times \left[{}_2F_1\left(-n+l+1, -n'+l, 2l, -\frac{4nn'}{(n-n')^2}\right)\right. \\ &\left. \max(l,l') - \left(\frac{n-n'}{n+n'}\right)^2 {}_2F_1\left(-n+l-1, -n'+l, 2l, -\frac{4nn'}{(n-n')^2}\right)\right] \end{aligned} \tag{B.1}$$

where $l' = l-1$. $M^2_{n,l\to n',l'}$ is symmetric, getting $l' = l+1$ for the same n and n' is just a matter of switching around indices. For heavier elements than hydrogen we would have to divide by the effective nuclear charge. *Gauss's hypergeometric functions* ${}_2F_1$ are series on the form:

$$
\begin{aligned}
{}_2F_1(a,b,c,x) &= 1 + \frac{ab}{c(1!)}x + \frac{a(a+1)b(b+1)}{c(c+1)(2!)}x^2 + \dots \\
&= \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{x^n}{n!}
\end{aligned}
\tag{B.2}
$$

with $(p)_n = \binom{p}{n} n!$ being the Pochhammer symbol which may be thought of as a rising factorial with n factors. (B.2) is finite if a or b is a non-positive integer, and infinite if c is a non-positive integer.

To compute the Einstein coefficients we therefore may compute the hypergeometric functions instead of the integral. Making use of the symmetry between a and b and the recursive relation from [2] (defining $F(a) = {}_2F_1(a,b,c,x)$):

$$
(a-c)F(a-1) = a(1-x)\left[F(a) - F(a+1)\right] + (a+bx-c)F(a)
$$

with initial conditions $F(0) = 1$, $F(-1) = 1 - \frac{b}{c}x$ we may compute a wide range of values of the function. Note that to compute Einstein coefficients for transitions near the Lyman limit, the hypergeometric function involves arithmetics with numbers which may even go beyond the 64-bit double precision range of $\gtrsim 10^{300}$.

B.2 Oscillator strength f_{osc} and Einstein coefficient A

Once we have the dipole moment the oscillator strength f_{osc} is simply:

$$
f_{\text{osc},n,l\to n',l'} = \left(\frac{1}{n^2} - \frac{1}{n'^2}\right) \frac{M^2_{n,l\to n',l'}}{3(2l+1)} \max(l,l'), \tag{B.3}
$$

assuming $n < n'$.

Combining (B.3) with (3.14) and the Einstein relations we arrive at an expression for A:

$$
A_{n,l\to n',l'} = \frac{64\pi^4 \nu^3_{n\to n'}}{3hc^3} \frac{\max(l,l')}{2l+1} e^2 a_0^2 M^2_{n,l\to n',l'} \tag{B.4}
$$

or alternatively expressed in terms of the angular momentum quantum numbers as in [45]:

$$
A_{n,l,j\to n',l',j'} = \frac{64\pi^4 \nu^3_{n\to n'}}{3hc^3} \begin{Bmatrix} l & j & 1/2 \\ j' & l' & 1 \end{Bmatrix}^2 (2j'+1) e^2 a_0^2 M^2_{n,l\to n',l'} \tag{B.5}
$$

where $\{\dots\}$ is the Wigner 6–j symbol (see [45] for details), e being the electron charge and a_0 the bohr radius. In my code we implemented both these approaches to verify the accuracy of the code.

From this we may compute the natural width of the a line γ, the decay probability P_{decay} and the recycle fraction f_{recycle}, in addition to the core optical depth τ_{core} and cross-section σ_{core}. Below are listed values of P_{if} and f_{recycle} for the first Lyman transitions:

n	$P_{nP\rightarrow 1S}$	f_{recycle}	n	$P_{nP\rightarrow 1S}$	f_{recycle}
2	1.0000	0.8817	14	0.7779	0.7770
3	0.8817	0.8390	15	0.7770	0.7761
4	0.8390	0.8177	16	0.7761	0.7754
5	0.8177	0.8053	17	0.7754	0.7748
6	0.8053	0.7972	18	0.7748	0.7743
7	0.7972	0.7917	19	0.7743	0.7738
8	0.7917	0.7877	20	0.7738	0.7734
9	0.7877	0.7847	21	0.7734	0.7731
10	0.7847	0.7824	22	0.7731	0.7728
11	0.7824	0.7806	23	0.7728	0.7725
12	0.7806	0.7791	24	0.7725	0.7722
13	0.7791	0.7779	25	0.7722	0.7720

Table B.1: Table of the first 24 computed decay probabilites to the ground state $P_{nP\rightarrow 1S}$ and the fraction of photons will cascade to the $2P$ state from a Lyn transition f_{recycle}. These quantites were computed earlier by [25,41]. f_{recycle} ignores decay to the ground state as the medium is assumed to be optically thick to this transition such that the photon is assumed to be absorbed immediately, causing a zero net effect.